

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 11:07:47 ; Search time 4744 Seconds  
(without alignments)  
10971.708 Million cell updates/sec

Title: US-10-069-541-5  
Perfect score: 1743  
Sequence: 1 atggctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	% Query					Description
	No.	Score	Match	Length	ID	
1	1473	84.5	1743	29	AY413298	AY413298 Homo sapi
2	1463	83.9	1743	29	AY413299	AY413299 Pan trogl
3	1375	78.9	4097	11	AK053063	AK053063 Mus muscu
4	1375	78.9	4306	11	AK034415	AK034415 Mus muscu
5	1156.2	66.3	1743	29	AY413300	AY413300 Mus muscu
6	518.8	29.8	672	29	AG157499	AG157499 Pan trogl
7	472.2	27.1	707	14	CD350164	CD350164 UI-M-FY0-
8	462.8	26.6	669	13	BY727598	BY727598 BY727598
9	404	23.2	516	10	BE233479	BE233479 139685 MA
10	329.8	18.9	650	10	BB626260	BB626260 BB626260
11	312.8	17.9	541	10	AW668962	AW668962 111664 MA
12	290	16.6	675	13	BY729567	BY729567 BY729567
13	274.2	15.7	524	10	BE723927	BE723927 198406 MA
14	263.2	15.1	800	9	AL669749	AL669749 AL669749
15	225.8	13.0	549	13	BW274870	BW274870 BW274870
16	212.4	12.2	1037	9	AL666817	AL666817 AL666817
17	210	12.0	941	14	CD360297	CD360297 AGENCOURT
18	209.2	12.0	641	12	BI630566	BI630566 RH59836.5
19	207.4	11.9	640	12	BI629504	BI629504 RH58381.5
20	205.2	11.8	658	12	BM629925	BM629925 170006875
21	203.2	11.7	652	10	BB626456	BB626456 BB626456
22	193.6	11.1	583	13	BW277281	BW277281 BW277281
23	192.6	11.0	605	13	BQ829470	BQ829470 LL6in2176
24	186.4	10.7	624	12	BJ122485	BJ122485 BJ122485
25	183	10.5	681	14	CD306544	CD306544 StrPu691.
26	178.6	10.2	565	12	BJ125564	BJ125564 BJ125564
27	177.4	10.2	576	14	CB391304	CB391304 OSTF149A8
28	176.6	10.1	310	9	AL918603	AL918603 AL918603
29	166	9.5	604	9	AU199794	AU199794 AU199794
30	163.4	9.4	500	9	AV994375	AV994375 AV994375
31	163.2	9.4	555	12	BJ117801	BJ117801 BJ117801
32	158.4	9.1	646	9	AB078155	AB078155 AB078155
c 33	157.4	9.0	801	13	BW002036	BW002036 BW002036
c 34	155	8.9	561	28	AQ316435	AQ316435 RPCI11-10
35	153	8.8	584	12	BJ105382	BJ105382 BJ105382
36	145.2	8.3	500	12	BP186503	BP186503 BP186503
c 37	134	7.7	632	28	AZ612750	AZ612750 1M0439J17
38	128.2	7.4	500	12	BJ105730	BJ105730 BJ105730
c 39	127.6	7.3	618	28	AZ908709	AZ908709 RPCI-24-2
40	121	6.9	926	29	CNS04L3J	AL295624 Tetraodon
c 41	117.2	6.7	525	12	BI508286	BI508286 BB170004A
c 42	117.2	6.7	530	12	BI503332	BI503332 BB170012A
43	112.8	6.5	355	9	AU209671	AU209671 AU209671
c 44	112.4	6.4	558	12	BI507950	BI507950 BB170010A
c 45	108	6.2	420	12	BI506529	BI506529 BB170027B

## ALIGNMENTS

## RESULT 1

AY413298

LOCUS AY413298 1743 bp DNA linear GSS 12-DEC-2003

DEFINITION Homo sapiens HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY413298

VERSION AY413298.1 GI:39769260

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .1743  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

gene <1. .>1743  
 /locus\_tag="HCM4844"

## ORIGIN

Query Match 84.5%; Score 1473; DB 29; Length 1743;

Best Local Similarity 84.5%; Pred. No. 0;

Matches 1473; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
 |||

Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
 |||

Db 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGNN	180
Qy	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	NN	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	NN	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	NN	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTCTGCACTGATGGGAGAAATGTTT	420
Db	361	NN	420
Qy	421	TGGGCTGCAGCAATTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	NN	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTGAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTGAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020



Db	961	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1080
Db	1021	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAG	1260
Db	1201	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	 CTGCTTTGTGTACTCTTTGTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	 TCTGGCCTCTTCTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTAAAA	1440
Db	1381	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTAAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	 TGA	1743

AY413299

LOCUS AY413299 1743 bp DNA linear GSS 12-DEC-2003

DEFINITION Pan troglodytes HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY413299

VERSION AY413299.1 GI:39769261

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Location/Qualifiers

source 1. .1743  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"

gene <1. .>1743  
/locus\_tag="HCM4844"

ORIGIN

Query Match 83.9%; Score 1463; DB 29; Length 1743;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 1466; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
|||||

Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
|||||

Db 61 GTTGAATATGGGCTGCCTGGAGAACCACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
|||||

Db 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACANN 180

Qy 181 ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

[illegible]

Qy	1081	CGGAACATCTACCAGCTTTCTCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCTCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACGTCAATTCTTAACCAACATTTGCGTCTCCTATCTAGCCAAATAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAAC TTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA	1680
Db	1621	GAAC TTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA	1680
Qy	1681	GAGGCC TTCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCC TTCTTGATGTTGATTCCAGTCCAGAANGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

# RESULT 3

AK053063

LOCUS AK053063 4097 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:D930038E20 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION AK053063

VERSION AK053063.1 GI:26343192  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
 AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE 2  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE 4  
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5  
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 4097)  
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
 Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .4097  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:D930038E20"  
/db\_xref="MGI:2424012"  
/db\_xref="taxon:10090"  
/clone="D930038E20"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="15 days embryo"

CDS 512. .2254  
/note="unnamed protein product; putative solute carrier family 5 (choline transporter), member 7 (MGD|MGI:1927126, GB|NM\_022025, evidence: BLASTN, 99%, match=1743)"  
/codon\_start=1  
/protein\_id="BAC35253.1"  
/db\_xref="GI:26343193"  
/translation="MSFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNPEERSEAIIV  
GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYGPCCGLAWAQAPIGYSLSLILGGLFF  
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD  
VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT  
AVHAKYQSPWLGTIESVEVYTWLDNFFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL  
AAFGCLVMALPAICIGAIGASTDWNQTAYGYDPKTKKEADMILPIVLQYLCVPYISF  
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVLVFGASA  
TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG  
EPYLYLQPLIFYPGYSDKNGIYNQRFPFKTLSMVTSFFTNICVSYLAKYLFESGTLF  
PKLDVFDVAVVARHSEENMDKTIIVRNENIKLNELAPVKPRQSLTSLSTFTNKEALLDV  
DSSPEGSGTEDNLQ"

ORIGIN

Query Match 78.9%; Score 1375; DB 11; Length 4097;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
||| ||||| || ||||| ||||| ||| | ||||| || || || |||  
Db 512 ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 571

Qy	61	GTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	572		631
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	632		691
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	692		751
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	752		811
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	812		871
Qy	361	ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	872		931
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	932		991
Qy	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	992		1051
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTACAGTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	1052		1111
Qy	601	ATCAGCGTCCCTTTTGCAATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	1112		1171
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	1172		1231
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	1232		1291
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Db	1292		1351
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1352		1411
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960

Db	1412	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	1471
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1472	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCCTGTT	1531
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1532	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGGGCGAGTTCTATGTTTGCT	1591
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1592	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1651
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1652	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1711
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCGTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1712	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG	1771
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1772	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1831
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1832	TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC	1891
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAA	1440
Db	1892	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAA	1951
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1952	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	2011
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	2012	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	2071
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAAATGAAAAATATTAAATTAGAT	1620
Db	2072	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAAATATCAAATTAAAT	2131
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	2132	GAACTTGCACTTGTGAAGCCACGAGAGCCTAACCCTCAGTTCAACTTTCACCAATAAG	2191
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	2192	GAGGCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA	2251
Qy	1741	TGA	1743



## RESULT 4

AK034415

LOCUS AK034415 4306 bp mRNA linear HTC 18-SEP-2003

DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330188K24 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION AK034415

VERSION AK034415.1 GI:26329926

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

## REFERENCE

2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

## REFERENCE

3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

## REFERENCE

4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

## REFERENCE

5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

## REFERENCE

6 (bases 1 to 4306)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,

Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/.

FEATURES

source

Location/Qualifiers  
1. .4306  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:9330188K24"  
/db\_xref="MGI:2398619"  
/db\_xref="taxon:10090"  
/clone="9330188K24"  
/sex="male"  
/tissue\_type="diencephalon"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"

CDS

394. .2136  
/note="unnamed protein product; putative  
solute carrier family 5 (choline transporter), member 7  
(MGD|MGI:1927126, GB|NM\_022025, evidence: BLASTN, 99%,  
match=1743)"  
/codon\_start=1  
/protein\_id="BAC28702.1"  
/db\_xref="GI:26329927"  
/translation="MSFHVEGLVAIILFYLLIFLVGIWAAWKTKNsgnPEERSEAIIV  
GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYGP GCGLAWAQAPIGYSLSLILGGLFF  
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD  
VNISVIVSALIAILYTLVGGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT  
AVHAKYQSPWLGTIESVEVYTWLDNFFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL  
AAFGLVMALPAICIGAIGASTDWNQTA YGYDPKTKEEADMILPIVLQYLCPVYISF  
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVLVFGASA  
TAMALLTKTVYGLWYLSDDLVIYIIIFPQLLCVLFIKGTNTYGAAGYIFGLFLRITGG  
EPYLYLQPLIFYPGYYS DKNGIYNQRFPFKTL SMVTSFFTNICVSYLAKYLFESGTLF  
PKLDVFDVAVVARHSEENMDKTI LVRNENIKLNELAPVKPRQSLTLSSTFTNKEALLDV

ORIGIN

[illegible]

Db 1114 CTTGATAATTTTCTGTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 1173

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840  
 || ||||| ||||| ||||| || ||||| ||||| ||||| |||||

Db 1174 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG 1233

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1234 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 1293

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1294 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC 1353

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1354 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1413

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1414 TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1473

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1474 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC 1533

Qy 1141 ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1534 ATGAGGATCACTGTGCTTGTGTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG 1593

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1594 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG 1653

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1654 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1713

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1714 TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC 1773

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1774 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1833

Qy 1441 ACACCTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1834 ACTCTCTCCATGGTTACCTCATTCCTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1893

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1894 CTATTTGAAAGTGGAACCTTGCCCTCCAAATTAGATGTATTTGATGCTGTTGTGCGCAAGG 1953

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1954 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 2013

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA 1680  
 ||||| ||||| || || ||||| | ||||| ||||| |||||  
 Db 2014 GAACTTGACCTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTACCAATAAG 2073  
 Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTACAG 1740  
 ||||| ||||| ||||| ||||| || ||||| ||||| ||||| |||||  
 Db 2074 GAGGCCCTCCTTGATGTTGATTCCAGTCCGAGGGTCTGGGACTGAAGATAACTTACAA 2133  
 Qy 1741 TGA 1743  
 |||  
 Db 2134 TGA 2136

# RESULT 5

AY413300

LOCUS AY413300 1743 bp DNA linear GSS 12-DEC-2003

DEFINITION Mus musculus HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY413300

VERSION AY413300.1 GI:39769262

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..1743  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 gene <1..>1743  
 /locus\_tag="HCM4844"

## ORIGIN

Query Match 66.3%; Score 1156.2; DB 29; Length 1743;  
 Best Local Similarity 73.1%; Pred. No. 3.8e-304;  
 Matches 1275; Conservative 0; Mismatches 468; Indels 0; Gaps 0;

Qy	1	ATGGCTTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCACAGTGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGNN	180
Qy	181	ACCTGGGTCGGAGGAGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	NN	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	NN	300
Qy	301	TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	NN	360
Qy	361	ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	NN	420
Qy	421	TGGGCTGCAGCAATTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	NN	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCGTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTCGGG	840

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||| ||||| | || ||||| ||| || || ||||| || |||||  
 Db 841 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATT'TTA 960  
 ||||| ||||| || ||| ||||| ||||| ||||| ||||| |||||  
 Db 901 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 961 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1021 TCAGCTGCTGTCTATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1081 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC 1140

Qy 1141 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||| ||||| ||||| ||||| ||||| |||||  
 Db 1141 ATGAGGATCACTGTGCTTGTGTTGCGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAG 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1201 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTTTACATCATCATCTTCCCACAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1261 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1321 TTTGGACTATTCTTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACCTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1441 ACTCTCTCCATGGTTACCTCATTTCTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1501 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTGCAAGG 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620

Qy 1621 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1621 GAACTTGCACCTGTGAAACCTCGGCAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAG 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

```

Db      1681  ||||| |||||||||||||||||||| || |||||||||||||||||||| |||||
          GAGGCCCTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA 1740

Qy      1741  TGA 1743
          |||
Db      1741  TGA 1743

```

## RESULT 6

```

AG157499
LOCUS          AG157499                672 bp    DNA        linear    GSS 09-JAN-2002
DEFINITION    Pan troglodytes DNA, clone: RP43-022H02.T7, genomic survey
               sequence.
ACCESSION     AG157499
VERSION       AG157499.1  GI:16687177
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
  ORGANISM    Pan troglodytes
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE     1
  AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
               Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE       BAC end sequences of Library RPCI-43
  JOURNAL     Unpublished
REFERENCE     2  (bases 1 to 672)
  AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
               Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE       Direct Submission
  JOURNAL     Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
               and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
               1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
               (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
               Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT       Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
               end was generated during the R&D process and may have higher chance
               of clone tracking errors.
               PRIMERS
                 Sequencing: T7
               LIBRARY
                 Vector      : pBACe3.6
                 R.Site 1   : EcoRI
                 R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
   source             1. .672
                       /organism="Pan troglodytes"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:9598"
                       /clone="RP43-022H02.T7"
                       /sex="male"
                       /cell_type="lymphocytes"
                       /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

```

Query Match 29.8%; Score 518.8; DB 29; Length 672;  
Best Local Similarity 97.3%; Pred. No. 3.5e-130;  
Matches 549; Conservative 0; Mismatches 12; Indels 3; Gaps 2;



Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTATGCGAATCACAGTGTGTTGGAGC 1169  
 | | ||||||||||||||||||||||||||||||||||||||||  
 Db 111 ACAGGCTTCGGACAAAGAAATCGTTTGGGTATGCGAATCACAGTGTGTTGGAGC 170

Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 1229  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 171 ATCTGCAACAGCCATGGCCTTGCTGACGAAGACTGTGTATGGGCTCTGGTACCTCAGTTC 230

Qy 1230 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 1289  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 231 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 290

Qy 1290 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG 1349  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 291 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG 350

Qy 1350 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 1409  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 351 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 410

Qy 1410 TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCCTTAAC 1469  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 411 TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACGTCATTCCTTAAC 470

Qy 1470 CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTG-AAAGTGGAACCTTGCCACCTA 1528  
 |||||||| |||||||||||| |||||||| ||||||||||||  
 Db 471 CAACATTTGCGTCTCCTATCTAGCCAAATATCTATTTGAAAAGTGGAACCTTGCCACCTA 530

Qy 1529 AATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAA 1588  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 531 AATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAA 590

Qy 1589 TTCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGA 1648  
 || |||||||||||||| |||||||| ||||||||||||  
 Db 591 TTTTGTCAAAAAATGAAA--TATAAATTAGATGACCTTGCACTTGTGAAGCCACGACAGA 648

Qy 1649 GCATGACCCCTCAGCTCAACTTTCA 1672  
 |||||||||||| ||||||||  
 Db 649 ACATGACCCCTCAGCTTAACCTTTCA 672

RESULT 7

CD350164

LOCUS CD350164 707 bp mRNA linear EST 09-JUL-2003

DEFINITION UI-M-FY0-cf1-h-10-0-UI.r1 NIH\_BMAP\_FY0 Mus musculus cDNA clone  
 IMAGE:6851099 5', mRNA sequence.

ACCESSION CD350164

VERSION CD350164.1 GI:31141679

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 707)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG	60
Db	27	ATGTCTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG	86
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	87	GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA	146
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGT'TTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	147	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGT'TTTACCATGACAGCC	206
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240



Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,  
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,  
 Kawaji,H., Kawasaki,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,  
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,  
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,  
 Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,  
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,  
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,  
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,  
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,  
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,  
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
 Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 MEDLINE 22354683  
 PUBMED 12466851  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,  
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,  
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,  
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,  
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,  
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,  
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
    source                   Location/Qualifiers  
                  1. .669  
                  /organism="Mus musculus"  
                  /mol\_type="mRNA"  
                  /db\_xref="taxon:10090"  
                  /clone="B730003H24"  
                  /tissue\_type="medulla oblongata"  
                  /dev\_stage="6 days neonate"  
                  /lab\_host="DH10B"  
                  /clone\_lib="RIKEN full-length enriched, 6 days neonate  
                  medulla oblongata"  
                  /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
                  prepared and sequenced in Mouse Genome Encyclopedia  
                  Project of Genome Exploration Research Group in Riken  
                  Genomic Sciences Center and Genome Science Laboratory in  
                  RIKEN. Division of Experimental Animal Research in Riken  
                  contributed to prepare mouse tissues. 1st strand cDNA was  
                  primed with a primer [5'  
                  GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was  
                  prepared by using trehalose thermo-activated reverse  
                  transcriptase and subsequently enriched for full-length by  
                  cap-trapper. cDNA went through one round of normalization  
                  to Rot = 20.0 and subtraction to Rot = 459.0. Second  
                  strand cDNA was prepared with the primer adapter of  
                  sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC  
                  3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
                  modified pBluescript KS(+) after bulk excision from Lambda  
                  FLC I."

#### ORIGIN

Query Match                   26.6%;   Score 462.8;   DB 13;   Length 669;  
Best Local Similarity       86.7%;   Pred. No. 6.8e-115;  
Matches 509;   Conservative       0;   Mismatches   78;   Indels       0;   Gaps       0;

```
Qy      1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTTCCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db      81 ATGTCTTTCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 140

Qy      61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db      141 GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 200

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||
Db      201 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 260

Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||
Db      261 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 320

Qy      241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db      321 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 380

Qy      301 TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
```



```

/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1PIG"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

# ORIGIN

```

Query Match          23.2%;  Score 404;  DB 10;  Length 516;
Best Local Similarity 86.4%;  Pred. No. 7.2e-99;
Matches 446;  Conservative 0;  Mismatches 70;  Indels 0;  Gaps 0;

```

```

Qy      1067 GTTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAG 1126
        |||| ||||||| | ||||||| |||| | ||||||||| |||| |||| |
Db      1    GTTCTATGTTTGTAGAAACATCTAGCAGCTCTCATTAGACAAAACGCTTCGGACAGGG 60

Qy      1127 AAATCGTTTGGGTTATGCGAATCACAGTGTGTGTTGGAGCATCTGCAACAGCCATGG 1186
        | ||||| ||||| ||||| ||||||| || ||||||| || ||||||||| |||||
Db      61    AGATCGTCTGGGTCATGCGGATCACAGTATTCGTGTTTGGTGCGTCTGCAACAGCCATGG 120

Qy      1187 CCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCG 1246
        ||||||||| || || ||||||||| ||||||||| || ||||| || |||||
Db      121    CCTTGCTGACCAAGACCGTGTATGGGCTCTGGTACCTCAGCTCCGACCTCGTCTACATCA 180

Qy      1247 TTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCG 1306
        ||||||||| ||||||| ||||| ||||| | ||||| ||||||| || |||||||
Db      181    TTATCTTCCCGCAGCTGCTCTGTGTGCTCTTCATCAAGGGGACCAACACGTACGGGGCCG 240

Qy      1307 TGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATC 1366
        ||||||| || || ||||||| ||||||| |||| || || ||||||| ||| |||
Db      241    TGGCAGGGTACATTGCTGGCCTTTTCCTGAGGGTAACCGGTGGAGAGCCATACCTGAACC 300

Qy      1367 TTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGA 1426
        | ||||||||| ||||||| ||||| ||| | ||||||||| |||||||
Db      301    TGCAGCCCTTGATCTTTTACCCTGGTATTACGTTGAAAAAATGGTATATATAATCAGA 360

Qy      1427 AATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCT 1486
        ||| ||||||| ||||||| || || || || ||||||| |||
Db      361    GATTCCCATTTAAACCCCTTGCCATGCTCACCTCCTTCTTATCCAACATTTGCATCTCTT 420

Qy      1487 ATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATG 1546
        ||||||||| ||||||||| ||||||||| ||||||||| ||||||| | |||||||
Db      421    ATCTAGCCAAATATCTATTTGAAAGTGGAACCTTGCCACCAAAATTAGATATGTTTGATG 480

Qy      1547 CTGTTGTTGCAAGACACAGTGAAGAAAACATGGATA 1582
        ||||||||| ||||||||| ||||||||| |||||||||
Db      481    CTGTTGTTGCAAGACACAGTGAAGAAAACATGGATA 516

```

## RESULT 10

BB626260

LOCUS BB626260 650 bp mRNA linear EST 26-OCT-2001

DEFINITION BB626260 RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9330170D24 5', mRNA sequence.

ACCESSION BB626260

VERSION BB626260.1 GI:16464298  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 650)  
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,  
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,  
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,  
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,  
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
 Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.  
 TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. . 10 (10), 1617-1630 (2000)  
 wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,  
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.  
 and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. .  
 10 (11), 1757-1771 (2000)  
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,  
 Sugahara,Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)  
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,  
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
 Hayashizaki,Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
 FEATURES Location/Qualifiers  
 source 1. .650  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="9330170D24"



```

/sex="male"
/tissue_type="diencephalon"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male
diencephalon"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:
BamHI"

```

# ORIGIN

```

Query Match          18.9%;  Score 329.8;  DB 10;  Length 650;
Best Local Similarity 86.0%;  Pred. No. 1.4e-78;
Matches 375;  Conservative 0;  Mismatches 60;  Indels 1;  Gaps 1;

```

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db      201 ATGTCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 260

Qy      61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db      261 GTTGAATATGGGCTGCATGGAACCAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 320

Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||||| ||||| || |||| || ||||| ||||| ||||| ||||| |||||
Db      321 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 380

Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      381 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 440

Qy      241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      441 GGTCTAGCTTNGGCTCANGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 500

Qy      301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| || |||| ||||| ||||| ||||| ||||| |||
Db      501 TTTTTTGC-GAACCTATGCNGTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG 559

Qy      361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420
      ||||| ||||| || || || || || |||| ||||| ||||| |||||
Db      560 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCCCTGATGGGAGAGATGTTT 619

```

Qy 421 TGGGCTGCAGCAATTT 436  
|||||  
Db 620 TGGGCTGCAGCCATTT 635

RESULT 11

AW668962

LOCUS AW668962 541 bp mRNA linear EST 25-APR-2001  
DEFINITION 111664 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION AW668962  
VERSION AW668962.1 GI:7525476  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 541)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,  
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,  
Chitko-McKown,C.G., Perteau,G., Holt,I., Karamycheva,S., Liang,F.,  
Quackenbush,J. and Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

PUBMED 11282978

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCCAGTCACGACG

Plate: 95 row: L column: 20

Seq primer: ATTTAGGTGACACTATAG.

FEATURES Location/Qualifiers

source

1. .541

/organism="Bos taurus"

/mol\_type="mRNA"

/db\_xref="taxon:9913"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="MARC 1BOV"

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from lymph node, ovary,  
fat, hypothalamus, and pituitary."

ORIGIN

Query Match 17.9%; Score 312.8; DB 10; Length 541;

Best Local Similarity 85.0%; Pred. No. 5.9e-74;  
Matches 350; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy      889 TCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCA 948
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      130 TCAACAGCCTGGAACCAGACTGCATACGGGCTCTTGCTCCCAGGGAGAAACAGGAGGCA 189

Qy      949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      190 GACATGATCTTGCCGATTGTCCTCAAGTATCTCTGCCCCGTGTACATTTCTTACTTTGGT 249

Qy      1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGT 1068
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      250 CTTGGAGCCGTTTCTGCTGCTGTCATGTCCTCAGCAGATTCTTCCATCTTGTCAGCAAGT 309

Qy      1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      310 TCGATGTTTGCCTCGCAACATCTACCAGCTTTCATTTCAGACAAAATGCTTCTGACAAGGAG 369

Qy      1129 ATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCC 1188
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      370 ATAGTCTGGGTCATGCGCATCACGGTATTTGTGTTTGGAGCTTCTGCGATGACCATGGCC 429

Qy      1189 TTGCTGACGAAAACACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTTACATCGTT 1248
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      430 TTGCTAACGAAGACGGTGTATGGGCTCTGGTACCTCAGCTCTGACCTGGTCTACATCATC 489

Qy      1249 ATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATG 1300
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      490 ATCTTCCCGCAGTTGCTCTGCGTGCTCTTCATCAAGGGTACCAACACGTATG 541
```

## RESULT 12

BY729567

LOCUS BY729567 675 bp mRNA linear EST 17-DEC-2002

DEFINITION BY729567 RIKEN full-length enriched, 12 days embryo spinal cord Mus musculus cDNA clone C530033E06 5', mRNA sequence.

ACCESSION BY729567

VERSION BY729567.1 GI:27142694

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 675)

AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,

Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,  
 Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,  
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,  
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,  
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,  
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,  
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,  
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
 Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,  
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,  
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,  
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,  
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,  
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,  
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

FEATURES Location/Qualifiers

source 1. .675



## RESULT 13

BE723927

LOCUS BE723927 524 bp mRNA linear EST 25-APR-2001

DEFINITION 198406 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE723927

VERSION BE723927.1 GI:10125223

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 524)

AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,  
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,  
Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F.,  
Quackenbush, J. and Keele, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

PUBMED 11282978

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG

Plate: 106 row: L column: 14

Seq primer: ATTTAGGTGACACTATAG.

## FEATURES

source Location/Qualifiers

1. .524

/organism="Bos taurus"

/mol\_type="mRNA"

/db\_xref="taxon:9913"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="MARC 4BOV"

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from day 20 and day 40  
embryos."

## ORIGIN

Query Match 15.7%; Score 274.2; DB 10; Length 524;

Best Local Similarity 84.7%; Pred. No. 2e-63;

Matches 331; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

Qy

889 TCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCA 948

```

      |||||  |||||
Db      130 TCAACATCCTGGAACCAGACTGCATACGGGCCCTCTTGCTCCCAGGGAGAAACAGGAGGCA 189
Qy      949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008
      |||||  ||  ||  |||||  ||  |||||
Db      190 GACATGATCTTGCCGATTGTCTCAAGTATCTCTGCCCCGTGTACATTTCTTACTTTGGT 249
Qy      1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTC-TTCCATCTTGTGCAAG 1067
      |||||  ||  |||||
Db      250 CTTGGAGCCGTTTCTGCTGCTGTCATGTCCTCAGCAGATTCTTTCCATCTTGTGCAAG 309
Qy      1068 TTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGA 1127
      |||  |||||  ||  |||||
Db      310 TTCGATGTTTGTCTGCAACATCTACCAGCTTTCATTAGACAAAATGCTTCTGACAAGGA 369
Qy      1128 AATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGC 1187
      ||  ||  |||||  |||||  |||||  ||  |||||
Db      370 GATAGTCTGGGTCATGCGCATCACGGTATTTGTGTTGGAGCTTCTGCGATGACCATGGC 429
Qy      1188 CTTGCTGACGAAAACGTGTGTATGGGCTC-TGGTACCTCAGTTCTGACCTTGTGTTACATCG 1246
      |||||  |||||  ||  |||||
Db      430 CTTGCTAACGAAGACGGTGTATGGGCTCTTGGTACCTCAGCTCTGACCTGGTCTACATCA 489
Qy      1247 TTATCTTCCCCCAGCTGCTTTGTGTACTCTT 1277
      |  |||||  |||  |||  ||  ||  |||||
Db      490 TCATCTTCCCGCAGTTGCTCTGCGTGCTCTT 520

```

# RESULT 14

AL669749

LOCUS AL669749 800 bp mRNA linear EST 14-JAN-2002

DEFINITION AL669749 directional larval cDNA library *Ciona intestinalis* cDNA clone 052ZB03 5', mRNA sequence.

ACCESSION AL669749

VERSION AL669749.1 GI:18143007

KEYWORDS EST.

SOURCE *Ciona intestinalis*

ORGANISM *Ciona intestinalis*

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; *Ciona*.

REFERENCE 1 (bases 1 to 800)

AUTHORS Genoscope.

TITLE *Ciona intestinalis* directional larval cDNA library

JOURNAL Unpublished (2002)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

IMPORTANT: this sequence may contain errors. The *Ciona intestinalis* library from which the clone was isolated may be contaminated with cDNAs from bacteria or other Eukarya.

Directional larval cDNA library originate from Dr.M.Branno, Stazione A.Dohrn, Naples, Italy, and was prepared in pBluescript2SK+.

FEATURES Location/Qualifiers

source 1..800

/organism="Ciona intestinalis"

```

/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="052ZB03"
/clone_lib="directional larval cDNA library"
/note="Vector: pBluescript2SK+"

```

ORIGIN

```

Query Match          15.1%;  Score 263.2;  DB 9;  Length 800;
Best Local Similarity 62.7%;  Pred. No. 2.3e-60;
Matches 504;  Conservative 0;  Mismatches 286;  Indels 14;  Gaps 6;

Qy      250 TGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGC- 308
      ||| | || |||| | |||| | | | | | | |||| |
Db      2   TGGACGCAAGCACCCATTGGATACGCTTGC GCGTTAATACTTGGCGGCTTATTCTTTGCG 61

Qy      309 AAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGG 368
      || | |||| | || | |||| | || | || | || | ||
Db      62 AAGTAAATGCGAAGTGAGGGATATGTGACGATGTTGGATCCACTGCAGCGCAACT-TGG 120

Qy      369 AAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTCTGGGCTGC 428
      | | |||| | | | | | |||| | |||| | |||
Db      121 TCGAGTAATGGGAGCGTTTCTTTATATACCTGCACTTGCTGGAGAATTATTCTGGTCTGC 180

Qy      429 AGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACAT 488
      || | || | | | |||| | || | || | || | || | || |
Db      181 AGCTATATTGGCCGCGTTGGGCGGTACCTTCA-TGTTATCATTGATCTTCATATAACTGC 239

Qy      489 TTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTC 548
      |||| | | | |||| | |||| | | |||| | || | || | ||
Db      240 AGCTGTAATAGTATCTGCATGCATTGCTGTTGTATACACCATGGCCGGTGGTCTTTACTC 299

Qy      549 TGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGGATCAGCGT 608
      || | || | | |||| | |||| | | |||| | || | || | |
Db      300 GGTGCTTATACAGATGTAGTTTCAGTTGATTTGCATATTCATTGGACTGTGGTTGAGCAT 359

Qy      609 CCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTGCATGCCAA 668
      || | || | | | |||| | || |||| | || | || | |
Db      360 TCCATTGCGGTTCACTCATCTGCTGTATCAGACATCGCCACTACAGCTTACCACTCACC 419

Qy      669 ATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGGCTTGATAG 728
      || | || | |||| | || | || | || | || | || |
Db      420 TAAC-----TGGCTTGGTACTTGGGATATTTGACCACTGGTCTATGGATCGACTC 470

Qy      729 TTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGGGTTCTCTC 788
      | | |||| | | | | | |||| | || |||| | || || | ||
Db      471 TGCTCTGCTACTGTTATTTGGTGGAAATACCGTGGCAAGTTTACTTTCAAAGAGTTTATC 530

Qy      789 TTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGGTGCCTGGT 848
      | | | | |||| | || | || | || | || | || | ||
Db      531 GNCNAAAAGNNCANGAAGCGCTCAGAAGCTTTCATTCATTGCTGCGTTCCGATGTTTGT 590

Qy      849 GATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAGAC 908
      ||| | || | | | || | || | || | || | || | || |
Db      591 CATGTCAATACCTTCGATATTNATCGGTGCAATTGCTGCATCTACAGATTGGGACGCAAC 650

Qy      909 TGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTACCAATTGT 968
      | | | || | || | | | | | | | | | || | || | || |

```



```

Db      651 ATCGTACGGCCTCCCAAGTCCAGTTGANAAAGGCGACCAAGCCAATATTCTACCCATTGT 710
Qy      969 TCTGCAGTATCTC-TGCCCTGTGTATATTTCTTTCTTT-GGTCTTGGTGCAGTTTCTGCT 1026
      || || || ||| ||||| | | || ||||| || ||||| || ||||| |||
Db      711 GCTTCAATACCTCACCCCTGTAGCTGTATCATTCTTTGGGGCTTGGCGCTGTTTCTGCT 770
Qy      1027 GCTGTTATGTCATCAGCAGATTCT 1050
      ||||| ||||| || || |||
Db      771 GCTGTNATGTCATCTGCCGACTCT 794

```

# RESULT 15

BW274870

LOCUS BW274870 549 bp mRNA linear EST 11-NOV-2002

DEFINITION BW274870 Nori Satoh unpublished cDNA library, gastrula and neurula  
Ciona intestinalis cDNA clone cign070c12 5', mRNA sequence.

ACCESSION BW274870

VERSION BW274870.1 GI:24855481

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 549)

AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.

TITLE Expressed genes in Ciona intestinalis (2002c)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

## FEATURES

source

Location/Qualifiers

1. .549

/organism="Ciona intestinalis"

/mol\_type="mRNA"

/db\_xref="taxon:7719"

/clone="cign070c12"

/tissue\_type="whole body"

/dev\_stage="gastrula and neurula"

/clone\_lib="Nori Satoh unpublished cDNA library, gastrula  
and neurula"

## ORIGIN

Query Match 13.0%; Score 225.8; DB 13; Length 549;

Best Local Similarity 64.4%; Pred. No. 3.4e-50;

Matches 354; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

```

Qy      11 ATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAATAT 70
      |||| | | | |||| |||| ||||| || ||| | | ||| |
Db      2 ATGTTCTGTTTGTAGTGNCTATTATCGTCTTCTACGTTGCTATTCTAGCGATCGGTATTT 61
Qy      71 GGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAG 130
      || || ||||| | | || | || || | | ||| |||| | |||| |
Db      62 ATGCAGCATGGAGGAAAAGAAGAACCGGAAGAGGAAACGAG---AGCGAGACAATCATGG 118

```

Qy 131 TTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCTACCTGGGTCG 190  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 119 TCGGGGGAAGAGACATCGGACTCTTTGTTGGAAGCTTTACTATGACTGCTACGTGGGTAG 178

Qy 191 GAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTT 250  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 179 GTGGTGGTTACATCAACGGCACAGCAGAAGTTGTATACACCCCGGGTTCCGGTCTACTGT 238

Qy 251 GGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAA 310  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 239 GGACACAAGCGCCATTTGGTTACGGCTGCAGCCTCATGCTTGGCGGGTTGTTTTTCGCTA 298

Qy 311 AACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAA 370  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 299 AGAAAATGCGGACTCAGGGTTACGTCACCATGCTGGATCCATTGCAACGTAAGCTTGGCA 358

Qy 371 AACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGCAG 430  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 359 GGCATGCGGGGGTCTGTTGTACTTACCAGCACTCTTGGGTGAAATATTCTGGTCAGCCG 418

Qy 431 CAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACATTT 490  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 419 CCATCCTTGCCGCTCTTGGCGGTACATTGTCCGTGATCATAGACCTTGATATTCGTATCT 478

Qy 491 CTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTG 550  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 479 CTGTCATTGTATCTGCATGTATTGCTGTGTTGTATACGTTGGTTGGTGGTCTGTATTCGG 538

Qy 551 TGGCCTACAC 560  
 | | | | | | | |  
 Db 539 TGGCTTATAC 548

Search completed: March 22, 2004, 15:16:56  
 Job time : 4765 secs

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 09:56:13 ; Search time 6973 Seconds  
(without alignments)  
10834.205 Million cell updates/sec

Title: US-10-069-541-5  
Perfect score: 1743  
Sequence: 1 atggcctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	8					Description
	No.	Score	Query Match Length	DB	ID	
	1	1743	100.0	1743	6	E49871 High-affini
	2	1743	100.0	1743	6	BD012719 High-affi
	3	1743	100.0	1743	9	AF276871 Homo sapi
	4	1743	100.0	1813	9	HSA401466 Homo sapi
	5	1743	100.0	5158	9	AB043997 Homo sapi
	6	1738.2	99.7	1743	6	AR268949 Sequence
	7	1394.2	80.0	1743	6	E49870 High-affini
	8	1394.2	80.0	1743	6	BD012718 High-affi
	9	1394.2	80.0	4904	10	AB030947 Rattus no
	10	1375	78.9	1743	10	AF276872 Mus muscu
	11	1373.4	78.8	1743	6	E49872 High-affini
	12	1373.4	78.8	1743	6	BD012720 High-affi
	13	1373.4	78.8	4938	6	AX080443 Sequence
	14	1367	78.4	1743	10	MMU401467 Mus muscu
	15	867	49.7	2528	5	TMA420808 Torpedo m
	16	730	41.9	1132	5	GGA511267 Gallus ga
	17	630.8	36.2	2239	9	HSA308384 Homo sapi
	18	630.8	36.2	190043	9	AC009963 Homo sapi
	19	502.8	28.8	232792	2	AC106657 Rattus no
c	20	501.2	28.8	155131	2	AC102873 Mus muscu
	21	431.4	24.8	3326	3	AY011119 Limulus p
	22	405.8	23.3	3255	3	AY047521 Drosophil
	23	363.8	20.9	1731	6	E49869 High-affini
	24	363.8	20.9	1731	6	BD012717 High-affi
	25	363.8	20.9	1985	3	AB030946 Caenorhab
	26	279.6	16.0	386	6	AX080449 Sequence
	27	242.6	13.9	1461	6	AX432086 Sequence
	28	226	13.0	1657	9	HSA308383 Homo sapi
	29	179.6	10.3	1178	9	HSA308378 Homo sapi
	30	179.6	10.3	186989	3	AC007812 Drosophil
	31	179.6	10.3	189117	3	AC009395 Drosophil
	32	179.6	10.3	255620	3	AE003723 Drosophil
c	33	167.6	9.6	140156	2	AC017381 Drosophil

	34	163	9.4	2326	9	HSA308379	AJ308379 Homo sapi
	35	155	8.9	1467	9	HSA308382	AJ308382 Homo sapi
c	36	151.8	8.7	40893	3	CBRG45E19	AC084631 Caenorhab
	37	150.2	8.6	736	9	HSA308381	AJ308381 Homo sapi
	38	150	8.6	1308	9	HSA308380	AJ308380 Homo sapi
c	39	141.6	8.1	39908	3	CEC48D1	Z81049 Caenorhabdi
c	40	141.6	8.1	330724	2	CEY67H2	AL022475 Caenorhab
c	41	132	7.6	152021	2	AC010923	AC010923 Drosophil
	42	98	5.6	616	11	G84799	G84799 S208P6036FB
c	43	51.8	3.0	10732	6	E32986	E32986 Gene encodi
c	44	49	2.8	2000	6	AX655393	AX655393 Sequence
	45	48.2	2.8	2781	1	BSU92466	U92466 Bacillus su

# ALIGNMENTS

## RESULT 1

E49871

LOCUS E49871 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49871

VERSION E49871.1 GI:22554902

KEYWORDS JP 2001136976-A/3.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: JP 2001136976-A 3 22-MAY-2001;

SCIENCE & TECH AGENCY

COMMENT OS Homo sapiens (human)

PN JP 2001136976-A/3

PD 22-MAY-2001

PF 27-DEC-1999 JP 1999368991

PI TATSUYA HAGA,TAKASHI OKUDA

PC C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,

PC C12N5/10,

PC C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC

FH Key Location/Qualifiers

FT CDS (1)..(1743).

FEATURES Location/Qualifiers

source 1..1743

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 1743; DB 6; Length 1743;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

|||||

Db 1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAACGCGATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAACGCGATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 |||  
 Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 |||  
 Db 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA 1080  
 |||  
 Db 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 |||  
 Db 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 |||  
 Db 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 |||  
 Db 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 |||  
 Db 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 |||  
 Db 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||  
 Db 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440

Qy 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||  
 Db 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||  
 Db 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||  
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 |||  
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||  
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

Qy 1741 TGA 1743

Db 1741 TGA 1743

RESULT 2  
BD012719  
LOCUS BD012719 1743 bp DNA linear PAT 02-AUG-2002  
DEFINITION High-affinity choline transporter.  
ACCESSION BD012719  
VERSION BD012719.1 GI:22092908  
KEYWORDS WO 0116315-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1743)  
AUTHORS Haga,T. and Okuda,T.  
TITLE High-affinity choline transporter  
JOURNAL Patent: WO 0116315-A 3 08-MAR-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP,TATSUYA HAGA,TAKASHI OKUDA  
COMMENT OS Homo sapiens (human)  
PN WO 0116315-A/3  
PD 08-MAR-2001  
PF 18-AUG-2000 WO 2000JP005545  
PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI  
TATSUYA HAGA,TAKASHI OKUDA  
PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC  
A61K38/17,  
PC A61K45/00,A61P25/28,G01N33/53,A01K67/027  
CC  
FH Key Location/Qualifiers  
FT CDS (1)..(1743).  
FEATURES Location/Qualifiers  
source 1..1743  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1743; DB 6; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
|  
Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
  
Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
|  
Db 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
  
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180  
|  
Db 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180  
  
Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
|



Db 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
 Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
 Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420  
 Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 Qy 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 Qy 601 ATCAGCGTCCCCCTTTGCATTGTGACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 ATCAGCGTCCCCCTTTGCATTGTGACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720  
 Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840  
 Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTTCCATGTTTGCA 1080  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTTCCATGTTTGCA 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 |||  
 Db 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140

Qy 1141 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 |||  
 Db 1141 ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 |||  
 Db 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 |||  
 Db 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320

Qy 1321 TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 |||  
 Db 1321 TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||  
 Db 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440

Qy 1441 ACACCTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||  
 Db 1441 ACACCTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||  
 Db 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||  
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 |||  
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||  
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

Qy 1741 TGA 1743  
 |||  
 Db 1741 TGA 1743

RESULT 3

AF276871

LOCUS AF276871 1743 bp mRNA linear PRI 27-NOV-2000

DEFINITION Homo sapiens high affinity choline transporter (SLC5A7) mRNA,  
 complete cds.

ACCESSION AF276871

VERSION AF276871.1 GI:10998441

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Apparsundaram,S., Ferguson,S.M., George,A.L. Jr. and Blakely,R.D.

TITLE Molecular cloning of a human, hemicholinium-3-sensitive choline transporter

JOURNAL Biochem. Biophys. Res. Commun. 276 (3), 862-867 (2000)

MEDLINE 20483599

PUBMED 11027560

REFERENCE 2 (bases 1 to 1743)

AUTHORS Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.

TITLE Direct Submission

JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA

FEATURES Location/Qualifiers

source 1. .1743  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2q12"

gene 1. .1743  
/gene="SLC5A7"

CDS 1. .1743  
/gene="SLC5A7"  
/note="hCHT; solute carrier family 5 member 7"  
/codon\_start=1  
/product="high affinity choline transporter"  
/protein\_id="AAG25940.1"  
/db\_xref="GI:10998442"  
/translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRTKNSGSAEERSEAIIV  
GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYVPGYGLAWAQAPIGYSLSLILGGLFF  
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD  
MHISVIIISALIATLYTLVGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT  
AVHAKYQKPWLGTVDSSSEVYSWLDSFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL  
AAGCLVMAIPAILIGAIGASTDWNQTAYGLPDKTTEADMILPIVLQYLCPVYISF  
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVFVFGASA  
TAMALLTKTVYGLWYLSSDLVYIVIFPQLLCVLFVKGTNTYGAVAGYVSGFLRITGG  
EPYLYLQPLIFYPGYPPDDNGIYNQKFPFKTLAMVTSFLTNICISYLAKYLFESGTLF  
PKLDVFDVAVARHSEENMDKTIILVKNENIKLDELALVKPRQSMTLSSTFTNKEAFLDV  
DSSPEGSGTEDNLQ"

# ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
|||||

Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
|||||

Db 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
 Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
 |||||  
 Db 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
 Qy 181 ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
 |||||  
 Db 181 ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
 Qy 241 GGCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGTGCGCTG 300  
 |||||  
 Db 241 GGCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGTGCGCTG 300  
 Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 |||||  
 Db 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420  
 |||||  
 Db 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420  
 Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 |||||  
 Db 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 Qy 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 |||||  
 Db 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 |||||  
 Db 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 Qy 601 ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 |||||  
 Db 601 ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 |||||  
 Db 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 |||||  
 Db 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840  
 |||||  
 Db 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840  
 Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 |||||  
 Db 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 |||||  
 Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 4  
 HSA401466  
 LOCUS HSA401466 1813 bp mRNA linear PRI 16-AUG-2000  
 DEFINITION Homo sapiens mRNA for high affinity choline transporter (CHT1 gene).  
 ACCESSION AJ401466  
 VERSION AJ401466.1 GI:9843753  
 KEYWORDS CHT1 gene; high affinity choline transporter.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Wieland,A., Bonisch,H. and Bruss,M.  
 TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1813)  
 AUTHORS Bruess,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-2000) Bruess M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
 FEATURES  
 source Location/Qualifiers  
 1. .1813  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2q11-13"  
 /tissue\_type="hypothalamus"  
 gene 1. .1813  
 /gene="CHT1"  
 CDS 19. .1761  
 /gene="CHT1"  
 /function="sodium- and chloride-dependent reuptake of choline"  
 /codon\_start=1  
 /evidence=experimental  
 /product="high affinity choline transporter"  
 /protein\_id="CAC03717.1"  
 /db\_xref="GI:9843754"  
 /db\_xref="GOA:Q9GZV3"  
 /db\_xref="SPTREMBL:Q9GZV3"  
 /translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRTKNSGSAEERSEAIIV  
 GGRDIGLLVGFTMTATWVGGGYINGTAEAVYVPGYGLAWAQAPIGYSLSLILGGLFF  
 AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAI FSALGATISVIIDVD  
 MHISVIIISALIATLYTLVGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT  
 AVHAKYQKPWLGTVDSSSEVSWLDSFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL  
 AAFGCLVMAIPAILIGAIGASTDWNQTAYGLPDKTTEADMILPIVLQYLCPVYISF  
 FGLGAVSAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIIVWMRITVFVFGASA  
 TAMALLTKTVYGLWYLSDDLVIYIVFPQLLCVLFVKGTNTYGAVAGYVSGFLFRITGG  
 EPYLYLQPLIFYPGYPPDDNGIYNQKFPFKTLAMVTSFLTNICISYLAKYLFESGTLF  
 PKLDVFDAVVARHSEENMDKTIILVKNENIKLDELALVKPRQSMTLSSTFTNKEAFLDV  
 DSSPEGSGTEDNLQ"  
 ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 1813;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      |||
Db     19 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 78

Qy     61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Db     79 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 138

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      |||
Db    139 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 198

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      |||
Db    199 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 258

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      |||
Db    259 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 318

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      |||
Db    319 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 378

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420
      |||
Db    379 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 438

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      |||
Db    439 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 498

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      |||
Db    499 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 558

Qy    541 CTCTATTCTGTGGCTTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      |||
Db    559 CTCTATTCTGTGGCTTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 618

Qy    601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACAGCTGTG 660
      |||
Db    619 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACAGCTGTG 678

Qy    661 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720
      |||
Db    679 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 738

Qy    721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
      |||
Db    739 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 798

Qy    781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
```

Db	799	 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGG	858
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	859	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	918
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	919	 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	978
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	979	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1038
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1080
Db	1039	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA	1098
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1099	 CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1158
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1159	 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1218
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1219	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1278
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1279	 CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1338
Qy	1321	TCTGGCCTCTTCTTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1339	 TCTGGCCTCTTCTTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1398
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1399	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1458
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1459	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1518
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1519	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1578
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1579	 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1638
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680



Db 1639 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1698

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||

Db 1699 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1758

Qy 1741 TGA 1743  
 |||

Db 1759 TGA 1761

# RESULT 5

AB043997

LOCUS AB043997 5158 bp mRNA linear PRI 19-NOV-2000

DEFINITION Homo sapiens mRNA for high-affinity choline transporter CHT1,  
 complete cds.

ACCESSION AB043997

VERSION AB043997.1 GI:11231080

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Okuda,T. and Haga,T.

TITLE Functional characterization of the human high-affinity choline  
 transporter

JOURNAL FEBS Lett. 484 (2), 92-97 (2000)

MEDLINE 20521663

PUBMED 11068039

REFERENCE 2 (bases 1 to 5158)

AUTHORS Okuda,T.

TITLE Direct Submission

JOURNAL Submitted (30-MAY-2000) Takashi Okuda, University of Tokyo, Faculty  
 of Medicine, Department of Neurochemistry; 7-3-1 Hongo, Bunkyo-ku,  
 Tokyo 1130033, Japan (E-mail:okuda@m.u-tokyo.ac.jp,  
 URL:http://park.ecc.u-tokyo.ac.jp/neurochemistry,  
 Tel:81-3-5841-3560, Fax:81-3-6814-8154)

FEATURES Location/Qualifiers

source

1. .5158

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="spinal cord"

CDS

277. .2019

/codon\_start=1

/product="high-affinity choline transporter CHT1"

/protein\_id="BAB18161.1"

/db\_xref="GI:11231081"

/translation="MAFHVEGLIAIIVFYLLILLVGIWAAWRKNSGSAEERSEAIIV  
 GGRDIGLLVGGFTMTATWVGGGYINGTAEAVYVPGYGLAWAQAPIGYSLSLILGGLFF  
 AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAFSALGATISVIIDVD  
 MHISVIIISALIATLYTLVGGLYSVAYTDVVQLFCIFVGLWISVPFALSHPAVADIGFT  
 AVHAKYQKPWLGTVDSSSEVYSWLDSFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL  
 AAFGCLVMAIPAILIGAIGASTDWNQTAYGLPDPKTTEEADMILPIVLQYLCPVYISF  
 FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVVWMRITVFVFGASA  
 TAMALLTKTVYGLWYLSDDLVIIVIFPQLLCVLFVKGTNTYGAVAGYVSGFLRITGG

EPYLYLQPLIFYPGYYPDDNGIYNQKFPEFKTLAMVTSFLTNICISYLAHYLFESGTLF  
 PKLDVFDVAVVHSEENMDKTIIVKNENIKLDELALVKPRQSMTLSSFTNKEAFLDV  
 DSSPEGSGTEDNLQ"

ORIGIN

Query Match 100.0%; Score 1743; DB 9; Length 5158;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	277	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	336
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	337	GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	396
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	397	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	456
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	457	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	516
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	517	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	576
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	577	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	636
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	637	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	696
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	697	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	756
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	757	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	816
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	817	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	876
Qy	601	ATCAGCGTCCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCCTGCTGTG	660
Db	877	ATCAGCGTCCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCCTGCTGTG	936
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	937	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	996

Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	997	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	1056
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	1057	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	1116
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1117	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	1176
Qy	901	AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1177	AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	1236
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTCTTTGGTCTTGGTGCAGTT	1020
Db	1237	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTCTTTGGTCTTGGTGCAGTT	1296
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCA	1080
Db	1297	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCA	1356
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1357	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1416
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1417	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1476
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1477	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1536
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1537	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1596
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1597	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1656
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1657	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1716
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1717	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1776
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1777	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1836
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620

```

      |||
Db      1837 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1896
      |||
Qy      1621 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAAA 1680
      |||
Db      1897 GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAAA 1956
      |||
Qy      1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
      |||
Db      1957 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 2016
      |||
Qy      1741 TGA 1743
      |||
Db      2017 TGA 2019

```

# RESULT 6

AR268949

LOCUS AR268949 1743 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 1 from patent US 6500643.

ACCESSION AR268949

VERSION AR268949.1 GI:29699686

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Wu,D.-H., Gu,Y., Millard,W.J. and He,Y.-J.

TITLE Human high affinity choline transporter

JOURNAL Patent: US 6500643-A 1 31-DEC-2002;

FEATURES Location/Qualifiers

source 1. .1743

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Query Match 99.7%; Score 1738.2; DB 6; Length 1743;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCACCTTCTAATTTTGCTG 60
      |||
Db      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCACCTTCTAATTTTGCTG 60
      |||
Qy      61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Db      61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      |||
Db      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      |||
Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      |||
Db      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      |||
Qy      241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

```

Db	241	 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACCTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACCTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTATCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTATCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140

Db 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 |||||  
 Db 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 |||||  
 Db 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 |||||  
 Db 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 |||||  
 Db 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||||  
 Db 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 Qy 1441 ACACCTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||||  
 Db 1441 ACACCTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||||  
 Db 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||||  
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA 1680  
 |||||  
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA 1680  
 Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||||  
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA 1740  
 Qy 1741 TGA 1743  
 |||  
 Db 1741 TGA 1743

# RESULT 7

E49870

LOCUS E49870 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49870

VERSION E49870.1 GI:22554901

KEYWORDS JP 2001136976-A/2.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

```

REFERENCE      1  (bases 1 to 1743)
AUTHORS       Haga,T. and Okuda,T.
TITLE         High-affinity choline transporter
JOURNAL       Patent: JP 2001136976-A 2 22-MAY-2001;
              SCIENCE & TECH AGENCY
COMMENT       OS    Rattus sp. (rat)
              PN    JP 2001136976-A/2
              PD    22-MAY-2001
              PF    27-DEC-1999 JP 1999368991
              PI    TATSUYA HAGA,TAKASHI OKUDA
              PC    C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,
              PC    C12N5/10,
              PC    C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC
              FH    Key                      Location/Qualifiers
              FT    CDS                      (1). .(1743).

FEATURES             Location/Qualifiers
     source           1. .1743
                     /organism="Rattus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10118"

```

## ORIGIN

Query Match 80.0%; Score 1394.2; DB 6; Length 1743;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGGCTG	60
Db	1	ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGTAATGCAGAAGAACGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480

Accession	Contig	Position	Sequence	Length
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	480	
Qy	481	ATGCACATTTCGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540	
Db	481	GTGAACATATCGGTTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCCTCGTGGGAGGG	540	
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600	
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	600	
Qy	601	ATCAGCGTCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTCACTGCTGTG	660	
Db	601	ATCAGTGTCCCATTTGCCCTGTGCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	660	
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720	
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTTGAATCAGTTGAAGTCTACACCTGG	720	
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780	
Db	721	CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	780	
Qy	781	GTTCTCTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTCGGG	840	
Db	781	GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCTTGGCAGCTTTTGGG	840	
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900	
Db	841	TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	900	
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960	
Db	901	AACCAAACCTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	960	
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGAGTT	1020	
Db	961	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTTCTTTGGGCTTGGTGCTGTT	1020	
Qy	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080	
Db	1021	TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1080	
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140	
Db	1081	CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140	
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200	
Db	1141	ATGAGGATCACTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1200	
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260	
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTCTACATCATCTTCCCACAG	1260	
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320	



Db 1261 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
| ||| || ||||| || ||||| ||||| || | ||||| |||

Db 1321 TTTGGACTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
||||||| ||||| | ||||| ||||| || |||||

Db 1381 TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
|| || ||||| ||||| ||||| | |||||

Db 1441 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
||||||| ||||| || ||||| ||||| |||||

Db 1501 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
||||||| ||||| ||||| ||||| ||||| || ||||| ||

Db 1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
||||||| || ||||| ||||| | ||||| |||||

Db 1621 GAACTTGCACTGTAAAGCCTCGACAGAGCCTAACCTCAGTTCAACTTTCACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
|||| ||||| ||||| || ||||| ||||| |||||

Db 1681 GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA 1740

Qy 1741 TGA 1743  
|||

Db 1741 TGA 1743

# RESULT 8

BD012718

LOCUS BD012718 1743 bp DNA linear PAT 02-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION BD012718

VERSION BD012718.1 GI:22092907

KEYWORDS WO 0116315-A/2.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: WO 0116315-A 2 08-MAR-2001;

JAPAN SCIENCE AND TECHNOLOGY CORP,TATSUYA HAGA,TAKASHI OKUDA

COMMENT OS Rattus norvegicus (rat)

PN WO 0116315-A/2

PD 08-MAR-2001

PF 18-AUG-2000 WO 2000JP005545

PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI

TATSUYA HAGA, TAKASHI OKUDA

PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC  
A61K38/17,

PC A61K45/00, A61P25/28, G01N33/53, A01K67/027

CC

FH Key Location/Qualifiers

FT CDS (1). .(1743).

FEATURES  
source Location/Qualifiers  
1. .1743  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"

#### ORIGIN

Query Match 80.0%; Score 1394.2; DB 6; Length 1743;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| ||||| ||||| ||| || ||| ||||| ||||| ||| ||| |||
Db      1 ATGCCTTTCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||| ||| ||||| ||||| |||||
Db     61 GTTGGAATATGGGCTGCATGGAACCAAAAACAGCGGTAATGCAGAAGAACGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| ||| ||| ||| ||| ||||| ||||| ||||| ||||| |||
Db    181 ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    301 TTTTGTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG 360

Qy    361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420
      ||||| ||||| ||| ||| ||||| ||| ||||| ||||| ||||| |||||
Db    361 ATCTATGGAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| ||||| ||| ||| ||| ||||| ||||| ||||| |||||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || |||| ||||| ||||| ||||| ||||| ||||| |||||
Db    481 GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      |||| ||||| || ||||| ||||| ||||| || ||||| ||||| |||||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG 600
```

Qy	601	ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCATTTGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTTCGGG	840
Db	781	GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCTTGGCAGCTTTTGGG	840
Qy	841	TGCCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCCTGGTGATGGCTCTACCAGCCATTTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAAACATGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTTTTCCCTGAGAATTACCGGAGGAGAGCCATATCTATACTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440

Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT	1620
Qy	1621	GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACCTGTAAAGCCTCGACAGAGCCTAACCCTCAGTTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCC'TTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

```

source      1. .4904
            /organism="Rattus norvegicus"
            /mol type="mRNA"

```

```

/strain="Wistar"
/db_xref="taxon:10116"
/clone="CHT1"
/tissue_type="spinal cord"
/clone_lib="rat spinal cord cDNA library"
/dev_stage="adult"
224. .1966
/codon_start=1
/product="high-affinity choline transporter CHT1"
/protein_id="BAA90484.1"
/db_xref="GI:6863034"
/translation="MPFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNAEERSEAIIV
GGRDIGLLVGGFMTATWVGGGYINGTAEAVYGP GCGLAWAQAPIGYSLSLILGGLFF
AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAI FSALGATISVIIDVD
VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT
AVHAKYQSPWLGTIESVEVYTWLDNLFLLMLGGIPWQAYFQRVLSSSSATYAQVLSFL
AAFGLVMALPAICIGAIGASTDWNQTAYGFDPDKTKEEADMILPIVLQYLCPVYISF
FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVFVFGASA
TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGA VAGYIFGLFLRITGG
EPYLYLQPLIFYPGYYPDKNGIYNQRFPFKTL SMVTSFFTNICVSYLAKYLFESGTL P
KLDIFDAVGSVRHSEENMDKTI LVRNENIKLNELAPVKPRQSLT LSTFTNK EALLDV
DSSPEGSSTDENLN"

```

Query Match		80.0%;	Score 1394.2;	DB 10;	Length 4904;
Best Local Similarity		87.5%;	Pred. No. 0;		
Matches 1525;		Conservative	0;	Mismatches	218; Indels 0; Gaps 0;
Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60		
Db	224	ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTTCTG	283		
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120		
Db	284	GTTGGAATATGGGCTGCATGGAACCAAAAACAGCGGTAATGCAGAAGAACGCAGCGAA	343		
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180		
Db	344	GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC	403		
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240		
Db	404	ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT	463		
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300		
Db	464	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG	523		
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360		
Db	524	TTTTTTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG	583		
Qy	361	ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420		
Db	584	ATCTATGGAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC	643		
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480		

Db	644		TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	703
Qy	481		ATGCACATTTCTGTCACTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	704		GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG	763
Qy	541		CTCTATTCTGTGGCCTACACTGATGTCGTTGAGCTCTTTTGCAATTTTGTAGGGCTGTGG	600
Db	764		CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	823
Qy	601		ATCAGCGTCCCTTTTGCAATGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	824		ATCAGTGTCCCATTTGCCCTGTGCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	883
Qy	661		CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	884		CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	943
Qy	721		CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	944		CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	1003
Qy	781		GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGTGCTCCTTCCTGGCAGCTTTCGGG	840
Db	1004		GTCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCCTGGCAGCTTTTGGG	1063
Qy	841		TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1064		TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG	1123
Qy	901		AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1124		AACCAAATGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	1183
Qy	961		CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1184		CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGTCTGTT	1243
Qy	1021		TCTGCTGCTGTTATGTCACTCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Db	1244		TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1303
Qy	1081		CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140
Db	1304		CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1363
Qy	1141		ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1364		ATGAGGATCACTGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1423
Qy	1201		ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTTACATCGTTATCTTCCCCAG	1260
Db	1424		ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1483
Qy	1261		CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320

Db 1484 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1543  
 Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| || ||||| ||||| || | ||||| ||  
 Db 1544 TTTGGACTTTTCCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1603  
 Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| | ||||| ||||| || ||||| ||  
 Db 1604 TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1663  
 Qy 1441 ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| | ||||| ||||| |||||  
 Db 1664 ACTCTCTCCATGGTTACCTCATTTCTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT 1723  
 Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||  
 Db 1724 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG 1783  
 Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| || ||||| ||  
 Db 1784 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAT 1843  
 Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 ||||| ||||| ||||| ||||| ||||| | ||||| ||||| ||||| |||||  
 Db 1844 GAACTTGCACTGTAAAGCCTCGACAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAA 1903  
 Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| ||||| ||||| ||||| ||||| || ||||| ||||| ||||| |||||  
 Db 1904 GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA 1963  
 Qy 1741 TGA 1743  
 |||  
 Db 1964 TGA 1966

# RESULT 10

AF276872

LOCUS AF276872 1743 bp mRNA linear ROD 28-FEB-2001

DEFINITION Mus musculus sodium and chloride-dependent high-affinity choline transporter mRNA, complete cds.

ACCESSION AF276872

VERSION AF276872.2 GI:13162669

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.

TITLE Molecular cloning and characterization of human and murine high-affinity choline transporters

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1743)

AUTHORS Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.

TITLE Direct Submission

JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at





Qy	301	TTCTTTTGC AAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCCGTTTCAGCAA	360
Db	301	TTTTTTTGC GAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTACAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCAGCTATTCTGCATTTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCTTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCTCTTCTTGGCAGCTTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTA CTGTCTCTTCTTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATTTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200

Db	1141	ATGAGGATCACTGTGCTTGTGTTCTGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTTCCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAT	1620
Qy	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA	1680
Db	1621	GAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGTTCAACTTTACCAATAAG	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCCTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAATTTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 11

E49872

LOCUS E49872 1743 bp DNA linear PAT 27-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION E49872

VERSION E49872.1 GI:22554903

KEYWORDS JP 2001136976-A/4.

SOURCE Mus sp.

ORGANISM Mus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE	High-affinity choline transporter					
JOURNAL	Patent: JP 2001136976-A 4 22-MAY-2001; SCIENCE & TECH AGENCY					
COMMENT	OS	Mus sp.	(mouse)			
	PN	JP 2001136976-A/4				
	PD	22-MAY-2001				
	PF	27-DEC-1999	JP 1999368991			
	PI	TATSUYA HAGA,TAKASHI OKUDA				
	PC	C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,				
	PC	C12N5/10,				
	PC	C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC				
	FH	Key	Location/Qualifiers			
	FT	CDS	(1). .(1743).			
FEATURES	Location/Qualifiers					
source	1. .1743 /organism="Mus sp." /mol_type="genomic DNA" /db_xref="taxon:10095"					
ORIGIN						
	Query Match	78.8%;		Score 1373.4;	DB 6;	Length 1743;
	Best Local Similarity	86.7%;		Pred. No. 0;		
	Matches 1512;	Conservative	0;	Mismatches 231;	Indels 0;	Gaps 0;
Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG	60			
Db	1	ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG	60			
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120			
Db	61	GTTGGAATATGGGCTGCATGGAAAACCACAAAACAGCGGCAACCCAGAAGAGCACAGTGAA	120			
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180			
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTGGTTTTACCATGACAGCC	180			
Qy	181	ACCTGGGTCCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240			
Db	181	ACCTGGGTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGATGGGCCAGGTTGT	240			
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300			
Db	241	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300			
Qy	301	TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360			
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG	360			
Qy	361	ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCTCTGCACTGATGGGAGAAATGTTTC	420			
Db	361	ATCTATGGAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC	420			
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480			
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480			
Qy	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540			

Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCTGTCACATCCTGCAGTCACCGACATCGGATTACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTGATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380

Db 1321 TTTGGACTATTTCCTGAGAACTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| | ||||| ||||| || |||||

Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| || || |||||

Db 1441 ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| || ||||| ||||| ||||| |||||

Db 1501 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

Db 1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 ||||| ||||| || || ||||| | ||||| ||||| |||||

Db 1621 GAACTTGCACTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| ||||| ||||| || ||||| ||||| |||||

Db 1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA 1740

Qy 1741 TGA 1743  
 |||

Db 1741 TGA 1743

# RESULT 12

BD012720

LOCUS BD012720 1743 bp DNA linear PAT 02-AUG-2002

DEFINITION High-affinity choline transporter.

ACCESSION BD012720

VERSION BD012720.1 GI:22092909

KEYWORDS WO 0116315-A/4.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)

AUTHORS Haga,T. and Okuda,T.

TITLE High-affinity choline transporter

JOURNAL Patent: WO 0116315-A 4 08-MAR-2001;

JAPAN SCIENCE AND TECHNOLOGY CORP,TATSUYA HAGA,TAKASHI OKUDA

COMMENT OS Mus musculus (mouse)

PN WO 0116315-A/4

PD 08-MAR-2001

PF 18-AUG-2000 WO 2000JP005545

PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI

TATSUYA HAGA,TAKASHI OKUDA

PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC  
 A61K38/17,

PC A61K45/00,A61P25/28,G01N33/53,A01K67/027

CC

FH Key Location/Qualifiers  
FT CDS (1)..(1743).

FEATURES Location/Qualifiers  
source 1..1743  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

ORIGIN

Query Match 78.8%; Score 1373.4; DB 6; Length 1743;  
Best Local Similarity 86.7%; Pred. No. 0;  
Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
||| ||||| || ||||| ||||| ||| | ||||| || || || |||  
Db 1 ATGTCTTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60  
  
Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||| |||  
Db 61 GTTGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120  
  
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||  
Db 121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGGTTGGTGGTTTACCATGACAGCC 180  
  
Qy 181 ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||  
Db 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 240  
  
Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||  
Db 241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300  
  
Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||  
Db 301 TTTTTCGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG 360  
  
Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420  
||||| ||||| || || || || || || ||||| ||||| ||||| |||||  
Db 361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT 420  
  
Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
||||| ||||| || || || ||||| ||||| ||||| ||||| ||||| |||  
Db 421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATGATGTGGAT 480  
  
Qy 481 ATGCACATTTCTGTCTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
|| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||  
Db 481 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCTAGTGGGTGGG 540  
  
Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGGCTGTGG 600  
||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||  
Db 541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600  
  
Qy 601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG 660  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 601 ATCAGTGTCCCTTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG 660

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 ||||| ||||| | || ||||| ||||| ||||| ||||| |||||  
 Db 661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG 720

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 721 CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTTGGCAGCTTTCGGG 840  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGTCTTCTTGGCAGCTTTTGGG 840

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 841 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 901 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 961 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCTCAGCAAGTTCCATGTTTGCA 1080  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1021 TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCTGGCGAGTTCTATGTTTGCT 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1081 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC 1140

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1141 ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1201 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1261 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320

Qy 1321 TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1321 TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1440

Qy 1441 ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1441 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

```

      |||
Db      1501 CTATTTGAAAGTGGAACCTTGCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1560
Qy      1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
      |||
Db      1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620
Qy      1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
      |||
Db      1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAG 1680
Qy      1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
      |||
Db      1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA 1740
Qy      1741 TGA 1743
      |||
Db      1741 TGA 1743

```

# RESULT 13

AX080443

LOCUS AX080443 4938 bp DNA linear PAT 22-FEB-2001

DEFINITION Sequence 1 from Patent WO0078950.

ACCESSION AX080443

VERSION AX080443.1 GI:13159872

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Sierzega,M. and Albrandt,K.

TITLE Differentially expressed genes in the adipocytes of obese mice

JOURNAL Patent: WO 0078950-A 1 28-DEC-2000;

AMYLIN PHARMACEUTICALS, INC. (US)

FEATURES

source

Location/Qualifiers

1. .4938

/organism="Mus musculus"

/mol\_type="unassigned DNA"

/db\_xref="taxon:10090"

/note="P4P6B1"

ORIGIN

Query Match 78.8%; Score 1373.4; DB 6; Length 4938;

Best Local Similarity 86.7%; Pred. No. 0;

Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      |||
Db      247 ATGTCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATACTTCTG 306
Qy      61 GTTGGAATATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Db      307 GTTGGAATATGGGCTGCATGGAAAACC AAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 366
Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180

```



Db	367	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC	426
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	427	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	486
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	487	GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	546
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	547	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCAACAG	606
Qy	361	ATCTATGGAACCGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	607	ATCTATGGAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT	666
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	667	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	726
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	727	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCTAGTGGGTGGG	786
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTGACGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	787	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	846
Qy	601	ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTGCTGTG	660
Db	847	ATCAGTGTCCCTTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG	906
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	907	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG	966
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	967	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	1026
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	1027	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGCTCCTTCCTGGCAGCTTTCGGG	1086
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1087	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	1146
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1147	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	1206
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020

Db 1207 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT 1266  
 Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCCATGTTTGCA 1080  
 || ||||| |||| |||| || || ||||| |||| || ||||| |||||  
 Db 1267 TCAGCTGCTGTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1326  
 Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1327 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC 1386  
 Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1387 ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG 1446  
 Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1447 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTCTACATCATCATCTTCCCACAG 1506  
 Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1507 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1566  
 Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1567 TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1626  
 Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1627 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1686  
 Qy 1441 ACACCTGCCATGGTTACATCATCTTAAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1687 ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1746  
 Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1747 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTGCGAAGG 1806  
 Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1807 CACAGTGAAGAGAACATGGACAAGACCATTTCTAGTCAGAAATGAAAATATCAAATTAAAT 1866  
 Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCCTCAGCTCAACTTTACCAATAAA 1680  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1867 GAACTTGCACTTGTGAACCTCGGCAGAGCCTAACCCCTCAGTTCAACTTTACCAATAAG 1926  
 Qy 1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1927 GAGGCCCTCCTTGATGTTGATTCCAGTCCGAGGGGTCTGGGACTGAAGATAACTTACAA 1986  
 Qy 1741 TGA 1743  
 |||  
 Db 1987 TGA 1989

LOCUS MMU401467 1743 bp mRNA linear ROD 16-AUG-2000  
 DEFINITION Mus musculus mRNA for high affinity choline transporter (CHT1 gene).  
 ACCESSION AJ401467  
 VERSION AJ401467.1 GI:9843808  
 KEYWORDS ChT1 gene; high affinity choline transporter.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Wieland,A., Bonisch,H. and Bruss,M.  
 TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1743)  
 AUTHORS Bruess,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-2000) Bruess M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
 FEATURES Location/Qualifiers  
 source 1. .1743  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/cJ"  
 /db\_xref="taxon:10090"  
 /tissue\_type="brainstem"  
 gene 1. .1743  
 /gene="CHT1"  
 CDS 1. .1743  
 /gene="CHT1"  
 /function="sodium- and chloride-dependent reuptake of choline"  
 /codon\_start=1  
 /evidence=experimental  
 /product="high affinity choline transporter"  
 /protein\_id="CAC03719.1"  
 /db\_xref="GI:9843809"  
 /db\_xref="GOA:Q9ESW5"  
 /db\_xref="SPTREMBL:Q9ESW5"  
 /translation="MSFHVEGLVAIILFYLLIFLVGIWAAWKTKNSGNPEEHSEAIIV  
 GGRDIGLLVGGFTMTATWVGGGYINGTAVAVYGPGLAWAQAPIGYSLSLILGGLFF  
 AKPMRSKGYVTMLDPFQQIYGKRMGGLLFIPALMGEMFWAAAIIFSALGATISVIIDVD  
 VNISVIVSALIAILYTLVGGLYSVAYTDVVQLFCIFIGLWISVPFALSHPAVTDIGFT  
 AVHAKYQSPWLGTIESVEVYTWLDNFLLLMLGGIPWQAYFQRVLSSSSATYAQVLSYL  
 AAFGCLVMALPAICIGAIGASTDWNQTAYGYDPKTKKEADMILPIVLQYLCVPYISF  
 FGLGAVSAAVMSSADSSILSASSMFARNIYQLSFRQNASDKEIVWVMRITVLVFGASA  
 TAMALLTKTVYGLWYLSSDLVYIIIFPQLLCVLFIKGTNTYGAVAGYIFGLFLRITGG  
 EPYLYLQPLIFYPGYYSKNGIYNQRFPFKTLMSMVTSFFTNICVSYLAKYLFESGTLF  
 PKLDVFDVAVVARHSEENMDKTIILVRNENIKLNELAPVKPRQSLTSLSTFTNKEALLDV  
 DSSPEGSGTEDNLQ"

# ORIGIN

Query Match 78.4%; Score 1367; DB 10; Length 1743;  
 Best Local Similarity 86.5%; Pred. No. 0;  
 Matches 1508; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
 ||| ||||| || ||||| ||||| ||| | ||||| || || || |||

Db 1 ATGTCCTTCCACGTAGAAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60

Qy 61 GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 61 GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
 ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||||

Db 121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 180

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
 ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGTAGCAGTGTATGGGCCAGGTTGT 240

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGATATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 301 TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTC AACAG 360

Qy 361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420  
 ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||

Db 361 ATCTATGGAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480

Qy 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 481 GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 540

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||

Db 541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy 601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 601 ATCAGTGTCCCTTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 660

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG 720

Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 721 CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy 781 GTTCTCTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG 840  
 || ||||| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||

Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGTCTACCTGGCAGCTTTTGGG 840

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900

Db	841	 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTGTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTCCCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGACCTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740

Db 1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA 1740

Qy 1741 TGA 1743

111

Db 1741 TGA 1743

# RESULT 15

TMA420808

LOCUS TMA420808 2528 bp mRNA linear VRT 27-NOV-2001

DEFINITION Torpedo marmorata mRNA for high affinity choline transporter (CHT1 gene).

ACCESSION AJ420808

VERSION AJ420808.1 GI:17148508

KEYWORDS CHT1 gene; high affinity choline transporter.

SOURCE Torpedo marmorata (marbled electric ray)

ORGANISM Torpedo marmorata

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea; Torpediniformes; Torpedinoidei; Torpedinidae; Torpedo.

REFERENCE 1

AUTHORS Guermonprez, L., O'Regan, S., Meunier, F.M. and Morot-Gaudry-Talarmain, Y.

TITLE Cyclosporin, FK506 and rapamycin inhibit neuronal choline uptake via calcineurin-dependent and independent mechanisms

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2528)

AUTHORS O'Regan, S.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2001) O'Regan S., Neurobiologie Cellulaire et Moleculaire, C.N.R.S., 1 av de la Terrasse, F-91198 Gif-sur-Yvette, FRANCE

FEATURES

Location/Qualifiers

source

1. .2528  
/organism="Torpedo marmorata"  
/mol\_type="mRNA"  
/db\_xref="taxon:7788"  
/clone="tH312"  
/tissue\_type="electric lobe"  
/tissue\_lib="lambda ZAPII ELL"

gene

1. .2528  
/gene="CHT1"

CDS

49. .1803  
/gene="CHT1"  
/function="neuronal Na-dependent choline transporter"  
/codon\_start=1  
/evidence=experimental  
/product="high affinity choline transporter"  
/protein\_id="CAD12727.1"  
/db\_xref="GI:17148509"  
/db\_xref="GOA:Q8UWF0"  
/db\_xref="SPTREMBL:Q8UWF0"

/translation="MTVHIDGIVAIVLFYLLILFVGLWAAWKSNTSMEGAMDRSEAI  
MIGGRDIGLLVGGFTMTATWVGGGYINGTAEAVYVPGYGLAWAQAPFGYALSIVIGGL  
FFAKPMRSRGYVTMLDPFQQMYGKRMGGLLFIPALLGEIFWSAAILSALGATLSVIVD  
ININVSVVVSAVIAVLYTLVGGLYSVAYTDVVQLFCIFLGLWISIPFALLNPAVDII  
VTANQEVYQEPWVGNIQSKDSLIIWIDNFLLLMLGGIPWQVYFQRVLSASSATYAQVLS

FLAAFGCVLMAIPSVLIGAIGTSTDWNQTSYGLPGPIGKNETDMILPIVLQHLCPPIYI  
SFFGLGAVSAAVMSSADSSILSASSMFARNIYHLAFRQEASDKEIVVWMRITIFLFGG  
AATSMALLAQSIYGLWYLSSDLVYVIIFPQLISVLFVKGTNTYGSYAGYIIGFLLRIS  
GGEPYLHMQPFIIYPGCYLDHSFGDDPVYVQRFPFKTMAMLFSLGNTGVSYLKYLEF  
VSGILPPKLDLDSVSVKHSKEIMDKTFLMNQDNITLSELVHVNPISASVSAALTNK  
EAFEDIEPNPELSKSGND"

polyA\_signal 2487. .2492  
/gene="CHT1"

ORIGIN

Query Match 49.7%; Score 867; DB 5; Length 2528;  
Best Local Similarity 69.3%; Pred. No. 2.9e-227;  
Matches 1217; Conservative 0; Mismatches 520; Indels 18; Gaps 2;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	49	ATGACCGTTCACATCGATGGGATCGTAGCGATCGTCCTGTTTACTTGTTAATCTTATTT	108
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCA-----GCGCAGAAGAGCGC	114
Db	109	GTTGGATTATGGGCTGCTTGGAAAAGTAAAAACACGTCAATGGAAGGAGCAATGGATCGG	168
Qy	115	AGCGAAGCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATG	174
Db	169	AGTGAAGCTATAATGATTGGGGGAAGAGATATCGGGCTGCTGTTGGTGGCTTCACAATG	228
Qy	175	ACAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCA	234
Db	229	ACCGCAACTTGGGTCGGTGGCGGTTATATCAATGGGACAGCAGAGGCGGTTTATGTTCT	288
Qy	235	GGTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTAGTCTGATTTTAGGT	294
Db	289	GGGTACGGCTTGGCCTGGGCGCAGGCTCCCTTCGGATACGCACTCAGCCTGGTTATTGGC	348
Qy	295	GGCCTGTTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTT	354
Db	349	GGCTTATTTTTCGCTAAACCCATGCGCTCACGGGGTTACGTGACCATGCTGGACCCGTTT	408
Qy	355	CAGCAAATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAA	414
Db	409	CAACAGATGTACGGTAAACGAATGGGAGGATTGCTCTTCATCCCCGCTCTCTGGGGGAA	468
Qy	415	ATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGAT	474
Db	469	ATCTTCTGGTCTGCAGCCATACTGTCCGCGCTAGGTGCAACTTTAAGCGTGATTGTGGAC	528
Qy	475	GTGGATATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTG	534
Db	529	ATCAATATAAACGTATCAGTGGTAGTTTCCGCTGTGATCGCTGTATTATACACTCTGGTC	588
Qy	535	GGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCAATTTTGTAGGG	594
Db	589	GGCGGGTTATACTCGGTCGCGTACACAGATGTCGTCCAGTTGTTTGCATCTTCTTAGGT	648
Qy	595	CTGTGGATCAGCGTCCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCCT	654
Db	649	TTGTGGATCAGCATCCCTTTTCGCCCTCTTAAATCCCGCTGTTACGGATATTATCGTGACC	708

Qy	655	GCTGTGCATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTAC	714
Db	709	GCAAATCAAGAAGTTTATCAGGAGCCTTGGGTGGGAATATACAATCAAAGGACAGTTTA	768
Qy	715	TCTTGGCTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTT	774
Db	769	ATCTGGATTGACAACCTTCTATTACTGATGCTGGGTGGAATCCCGTGGCAAGTATATTTT	828
Qy	775	CAGAGGGTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTTGGCAGCT	834
Db	829	CAGAGAGTCCTTTCTGCTTCTTCTGCTACCTATGCGCAAGTCCTGTCCTTTCTGGCTGCC	888
Qy	835	TTCGGGTGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACA	894
Db	889	TTCGGTTGCGTTCTCATGGCCATCCCGTCTGTTCTCATCGGTGCAATAGGAACATCCACT	948
Qy	895	GACTGGAACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATG	954
Db	949	GACTGGAATCAGACTTCCTATGGCTTGCCAGGCCCTATAGGCAAAAATGAGACTGATATG	1008
Qy	955	ATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGT	1014
Db	1009	ATTTTGGCGATCGTGCTGCAGCATCTGTGTCCACCCTACATTTCTTTTTTGGTCTTGGC	1068
Qy	1015	GCAGTTTCTGCTGCTGTTATGTATCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATG	1074
Db	1069	GCTGTCTCTGCTGCTGTGATGTATCATCGGCTGATTCTTCTATCTTATCAGCAAGTCTATG	1128
Qy	1075	TTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTT	1134
Db	1129	TTTGTCTCGGAATATTTACCATCTTGCTTTTCAGACAAGAGGCTTCAGACAAAGAAATAGTG	1188
Qy	1135	TGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTG	1194
Db	1189	TGGGTAATGCGAATCACCATATTTCTATTTGGAGGAGCTGCAACATCTATGGCATTGCTT	1248
Qy	1195	ACGAAAACGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTC	1254
Db	1249	GCTCAATCAATCTATGGCCTCTGGTATCTGAGCTCAGATCTTGTCTACGTCATTATCTTT	1308
Qy	1255	CCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGT	1314
Db	1309	CCTCAATTAATATCAGTGCTCTTCGTCAAGGGAACAAACACATATGGGTCTATTGCTGGA	1368
Qy	1315	TATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCC	1374
Db	1369	TATATCATTGGCTTTTTGCTTCGGATTAGTGGTGGTGAACCATATTTACATATGCAGCCA	1428
Qy	1375	TTGATCTTCTACCCTGGCT-----ATTACCCTGATGATAATGGTATATATAAT	1422
Db	1429	TTTATTTATTACCCTGGATGCTATTTAGATCATTCCTTTGGAGATGATCCTGTTTATGTT	1488
Qy	1423	CAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATC	1482
Db	1489	CAGAGATTTCCCTTTAAACCATGGCAATGTTATTCTCCTTCTTGGGCAACACTGGTGTA	1548



Qy 1483 TCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTT 1542  
 || |||| | ||||| || || | ||||| ||||| ||||| | ||  
 Db 1549 TCATATCTTGTCAAGTACCTGTTTCGTAAGTGGAATATTGCCACCAAATTAGACTTCCTT 1608

Qy 1543 GATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAAT 1602  
 || |||| | ||||| | |||| | |||| | || | | || |  
 Db 1609 GACAGCGTTGTATCAAAACACAGTAAGGAAATCATGGACAAAACATTCTTGATGAATCAG 1668

Qy 1603 GAAAATATTAAATTAGATGAACCTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGC 1662  
 || ||||| || || || | || || || || || | | ||  
 Db 1669 GACAATATTACTTTGTCAGAGCTGGTGCATGTTAATCCAATACACAGTGCTTCAGTTAGT 1728

Qy 1663 TCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGG 1722  
 | |||| ||||| || || || || |||| | | || || || || |  
 Db 1729 GCTGCTTTGACCAATAAGGAAGCATTTGAAGACATTGAGCCAAATCCTGAACTTTCTAAG 1788

Qy 1723 ACTGAAGATAATTTA 1737  
 | | || || |  
 Db 1789 TCAGGCAATGATTGA 1803

Search completed: March 22, 2004, 13:57:28  
 Job time : 6988 secs

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 09:53:04 ; Search time 742 Seconds  
 (without alignments)  
 9979.258 Million cell updates/sec

Title: US-10-069-541-5  
 Perfect score: 1743  
 Sequence: 1 atggcctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
 1: geneseqn1980s:\*  
 2: geneseqn1990s:\*  
 3: geneseqn2000s:\*  
 4: geneseqn2001as:\*  
 5: geneseqn2001bs:\*  
 6: geneseqn2002s:\*  
 7: geneseqn2003as:\*  
 8: geneseqn2003bs:\*  
 9: geneseqn2003cs:\*  
 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query				Description
		Match	Length	DB	ID	
1	1743	100.0	1743	4	AAF81712	Aaf81712 Human hig
2	1743	100.0	1743	5	AAH49207	Aah49207 Human CHO
3	1743	100.0	1743	9	ADD50638	Add50638 cDNA enco
4	1743	100.0	1813	9	ADD50646	Add50646 High-affi
5	1738.2	99.7	1743	8	ABX94338	Abx94338 Human cDN
6	1394.2	80.0	1743	4	AAF81711	Aaf81711 Rat high
7	1394.2	80.0	4904	9	ADD50642	Add50642 cDNA enco

8	1375	78.9	1743	9	ADD50640	Add50640 cDNA enco
9	1375	78.9	1743	9	ADD50660	Add50660 cDNA enco
10	1373.4	78.8	1743	4	AAF81713	Aaf81713 Mouse hig
11	1373.4	78.8	4938	5	AAD02457	Aad02457 Mouse P4P
12	630.8	36.2	8760	5	AAH49206	Aah49206 Human CHO
13	630.8	36.2	119040	9	ADD50656	Add50656 BAC seque
14	630.8	36.2	142299	9	ADD50651	Add50651 BAC seque
15	363.8	20.9	1731	4	AAF81710	Aaf81710 C. elegan
16	363.8	20.9	1985	9	ADD50644	Add50644 High-affi
17	279.6	16.0	386	5	AAD02461	Aad02461 Mouse P4P
18	266.6	15.3	1729	4	ABL29569	Abl29569 Drosophil
19	242.6	13.9	1461	6	ABK73210	Abk73210 Bacillus
20	180.8	10.4	10140	5	AAH49201	Aah49201 Human CHO
c 21	180.8	10.4	119040	9	ADD50656	Add50656 BAC seque
c 22	180.8	10.4	142299	9	ADD50651	Add50651 BAC seque
c 23	179.6	10.3	4223	4	ABL29568	Abl29568 Drosophil
24	167.6	9.6	1094	5	AAS82193	Aas82193 DNA encod
25	163	9.4	240	5	AAH49203	Aah49203 Human CHO
26	155	8.9	455	4	AAI11913	Aai11913 Probe #18
27	155	8.9	455	4	ABA53620	Aba53620 Human foe
28	155	8.9	455	4	AAI33245	Aai33245 Probe #19
29	155	8.9	455	4	ABA23372	Aba23372 Probe #18
30	155	8.9	455	4	AAK27341	Aak27341 Human bon
31	155	8.9	455	4	AAK01883	Aak01883 Human bra
32	155	8.9	455	4	ABS26913	Abs26913 Human liv
33	155	8.9	455	5	AAI01851	Aai01851 Probe #18
34	154.6	8.9	300	5	AAH49204	Aah49204 Human CHO
35	150.8	8.7	240	5	AAH49205	Aah49205 Human CHO
36	108	6.2	180	5	AAH49202	Aah49202 Human CHO
37	72	4.1	96	4	AAI21121	Aai21121 Probe #11
38	72	4.1	96	4	ABA66205	Aba66205 Human foe
39	72	4.1	96	4	AAI46388	Aai46388 Probe #15
40	72	4.1	96	4	ABA33269	Aba33269 Probe #11
41	72	4.1	96	4	AAK40366	Aak40366 Human bon
42	72	4.1	96	4	AAK14621	Aak14621 Human bra
43	72	4.1	96	4	ABS39934	Abs39934 Human liv
44	72	4.1	96	5	AAI06841	Aai06841 Probe #68
45	60	3.4	60	6	ABN37501	Abn37501 Human spl

#### ALIGNMENTS

##### RESULT 1

AAF81712

ID AAF81712 standard; cDNA; 1743 BP.

XX

AC AAF81712;

XX

DT 01-JUN-2001 (first entry)

XX

DE Human high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;  
KW ss.

XX

OS Homo sapiens.



Db	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTGCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080

Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

# RESULT 2

AAH49207

ID AAH49207 standard; cDNA; 1743 BP.

XX

AC AAH49207;

XX

DT 26-NOV-2001 (first entry)

XX  
DE Human CHOT encoding cDNA.  
XX  
KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;  
KW neuroprotective; gene therapy; antisense therapy; degenerative disease;  
KW cognitive disorder; Alzheimer's disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN DE10009055-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 28-FEB-2000; 2000DE-01009055.  
XX  
PR 28-FEB-2000; 2000DE-01009055.  
XX  
PA (BRUE/) BRUESS M.  
PA (BOEN/) BOENISCH H.  
XX  
PI Bruess M, Boenisch H;  
XX  
DR WPI; 2001-590709/67.  
DR P-PSDB; AAB86837.  
XX  
PT A new gene encoding human choline transporter, designated hCHOT is  
PT located on chromosome 2q11-13 and is useful to treat degenerative  
PT disorders such as Alzheimer's disease.  
XX  
PS Disclosure; Page 11; 12pp; German.  
XX  
CC This invention describes a novel gene encoding human choline transporter,  
CC designated hCHOT which is located on chromosome 2q11-13. The products of  
CC the invention have nootropic and neuroprotective activity and can be used  
CC for gene or antisense therapy. (I) is used to treat degenerative disease,  
CC particularly cognitive disorders such as Alzheimer's disease. Sense and  
CC antisense oligonucleotides derived from the gene may be used in  
CC diagnostics and other techniques. This sequence encodes the human CHOT  
CC protein described in the invention  
XX  
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 5; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTTCCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCATGTGGAAGGACTGATAGCTATCATCGTGTTCCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180

Qy	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAAGTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAAGTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTCTTTCTTTGGTCTTGGTGCAGTT	1020



Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAGTTCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Qy	1621	GAACTTGCACTTGTTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACTTGTTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743

RESULT 3

ADD50638

ID ADD50638 standard; cDNA; 1743 BP.

XX

AC ADD50638;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE cDNA encoding human high-affinity choline transporter (hCHT).  
 XX  
 KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;  
 KW cholinergic function; Parkinson's disease; Huntington's disease;  
 KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;  
 KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;  
 KW nootropic; neuroprotective; neuroleptic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1. .1743  
 FT /\*tag= a  
 FT /product= "hCHT"  
 XX  
 PN US2003114399-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 23-JUL-2001; 2001US-00911077.  
 XX  
 PR 23-JUL-2001; 2001US-00911077.  
 XX  
 PA (BLAK/) BLAKELY R D.  
 PA (APPA/) APPARSUNDARAM S.  
 PA (FERG/) FERGUSON S.  
 XX  
 PI Blakely RD, Apparsundaram S, Ferguson S;  
 XX  
 DR WPI; 2003-810914/76.  
 DR P-PSDB; ADD50639.  
 XX  
 PT Novel isolated polynucleotide encoding human or mouse high affinity  
 PT choline transporter polypeptide, useful in gene therapy to increase  
 PT cholinergic function in a cell of a patient suffering from Alzheimer's  
 PT disease.  
 XX  
 PS Claim 2; SEQ ID NO 1; 74pp; English.  
 XX  
 CC The present invention relates to the isolation of polynucleotide  
 CC sequences encoding human and mouse high-affinity choline transporter  
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
 CC polynucleotide sequence when delivered to a cell, increases cholinergic  
 CC function in the cell that is in a patient having Parkinson's disease,  
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
 CC myasthenia gravis. The hCHT antibody is useful for controlling  
 CC transporter CHT proteins to the brain, and for treating the above  
 CC mentioned diseases. The antibody is also useful for diagnosing the above  
 CC mentioned disorders and to detect the influence of cholinergic  
 CC signalling. The present sequence encodes hCHT. Note: The sequence data  
 CC for this patent was obtained in electronic format directly from the USPTO

CC web site at seqdata.uspto.gov.

XX

SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 9; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60
      |||
Db      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60

Qy     61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      |||
Db     61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      |||
Db    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      |||
Db    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      |||
Db    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300

Qy    301 TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      |||
Db    301 TTCTTTGCAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360

Qy    361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420
      |||
Db    361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      |||
Db    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      |||
Db    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      |||
Db    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660
      |||
Db    601 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG 660

Qy    661 CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720
      |||
Db    661 CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720

Qy    721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGG 780
```

Db	721	 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTGGTCTTGGTGCAGTT	1020
Db	961	 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTGGTCTTGGTGCAGTT	1020
Qy	1021	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Qy	1261	 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Qy	1561	 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620

Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA 1680  
 |||

Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||

Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740

Qy 1741 TGA 1743  
 |||

Db 1741 TGA 1743

#### RESULT 4

ADD50646

ID ADD50646 standard; DNA; 1813 BP.

XX

AC ADD50646;

XX

DT 15-JAN-2004 (first entry)

XX

DE High-affinity choline transporter (CHT) associated DNA sequence #2.

XX

KW High-affinity choline transporter; CHT; cholinergic function;  
 KW Parkinson's disease; Huntington's disease; Alzheimer's disease;  
 KW schizophrenia; dysautonomia; myasthenia gravis; brain;  
 KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
 KW neuroprotective; neuroleptic; ds.

XX

OS Unidentified.

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX

PF 23-JUL-2001; 2001US-00911077.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (BLAK/) BLAKELY R D.

PA (APPA/) APPARSUNDARAM S.

PA (FERG/) FERGUSON S.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX

DR WPI; 2003-810914/76.

XX

PT Novel isolated polynucleotide encoding human or mouse high affinity  
 PT choline transporter polypeptide, useful in gene therapy to increase  
 PT cholinergic function in a cell of a patient suffering from Alzheimer's  
 PT disease.

XX

PS Disclosure; SEQ ID NO 9; 74pp; English.

XX

CC The present invention relates to the isolation of polynucleotide

CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic  
CC function in the cell that is in a patient having Parkinson's disease,  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signalling. The present DNA sequence of unknown function is provided in  
CC the electronic sequence data but is not mentioned in the printed  
CC specification. Note: The sequence data for this patent was obtained in  
CC electronic format directly from the USPTO web site at seqdata.uspto.gov.  
XX  
SQ Sequence 1813 BP; 440 A; 406 C; 417 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 9; Length 1813;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
|  
Db 19 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 78  
  
Qy 61 GTTGAATATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
|  
Db 79 GTTGAATATGGGCTGCCTGGAGAACC AAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 138  
  
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180  
|  
Db 139 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 198  
  
Qy 181 ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
|  
Db 199 ACCTGGGTTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 258  
  
Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
|  
Db 259 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 318  
  
Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
|  
Db 319 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 378  
  
Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 420  
|  
Db 379 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC 438  
  
Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
|  
Db 439 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 498  
  
Qy 481 ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
|  
Db 499 ATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 558

Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	559	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	618
Qy	601	ATCAGCGTCCCCCTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	619	ATCAGCGTCCCCCTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	678
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	679	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	738
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	739	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	798
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	799	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	858
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	859	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	918
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	919	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	978
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	979	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1038
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Db	1039	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1098
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1099	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1158
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1159	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1218
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1219	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1278
Qy	1261	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1279	CTGCTTTGTGTACTCTTTGTAAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1338
Qy	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1339	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1398

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||  
 Db 1399 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1458

Qy 1441 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||  
 Db 1459 ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1518

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||  
 Db 1519 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1578

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||  
 Db 1579 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1638

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 |||  
 Db 1639 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1698

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||  
 Db 1699 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1758

Qy 1741 TGA 1743  
 |||  
 Db 1759 TGA 1761

RESULT 5  
 ABX94338

ID ABX94338 standard; cDNA; 1743 BP.

XX

AC ABX94338;

XX

DT 13-JUN-2003 (first entry)

XX

DE Human cDNA encoding high affinity choline transporter, HACT.

XX

KW Human; ss; gene; HACT; high affinity choline transporter; pain;  
 KW neurotransmitter biosynthesis; learning and memory; aging; epilepsy;  
 KW neurological disorder; spasticity; myoclonus; muscle spasm;  
 KW muscle hyperactivity; stroke; head trauma; neuronal cell death;  
 KW multiple sclerosis; spinal chord injury; dystonia; Alzheimer's disease;  
 KW Myasthenia Gravis; multi-infarct dementia; AIDS dementia;  
 KW Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;  
 KW ALS; attention deficit disorder; organic brain syndrome; schizophrenia;  
 KW nicotine addiction; memory disorder; cognitive disorder.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	1. .1743
FT		/*tag= a
FT		/product= "HACT"

XX

PN US6500643-B1.



XX  
 PD 31-DEC-2002.  
 XX  
 PF 07-SEP-2000; 2000US-00657252.  
 XX  
 PR 07-SEP-2000; 2000US-00657252.  
 XX  
 PA (UYFL ) UNIV FLORIDA.  
 XX  
 PI Wu D, Gu Y, Millard WJ, He Y;  
 XX  
 DR WPI; 2003-361535/34.  
 DR P-PSDB; ABU08979.  
 XX  
 PT Novel isolated polynucleotide (I) that encodes high affinity choline  
 PT transporter protein, useful for preventing, treating or ameliorating  
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's  
 PT disease.  
 XX  
 PS Claim 2; Col 17-21; 20pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide which encodes a high  
 CC affinity choline transporter (HACT) protein appearing as ABU08979. Also  
 CC included are a polynucleotide encoding a fragment consisting of at least  
 CC about 50 amino acids of the HACT protein, a vector comprising the  
 CC polynucleotide, a composition comprising a vector comprising a  
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids  
 CC of a polynucleotide appearing as ABX94339 (encoding choline  
 CC acetyltransferase), a recombinant host cell which comprises the vector  
 CC (used to express the HACT protein or fragment). The polynucleotide is  
 CC useful as a probe or primer to detect the presence of HACT polynucleotide  
 CC in a sample, such as a biological sample, or for screening for test  
 CC agents which bind to the polynucleotide. A pharmaceutical composition  
 CC comprising the polynucleotide is useful for preventing, treating or  
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,  
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head  
 CC trauma, neuronal cell death, multiple sclerosis, spinal chord injury,  
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi- infarct  
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,  
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine  
 CC addiction, organic brain syndromes, schizophrenia or memory and cognitive  
 CC disorders. HACT is thought to be the rate limiting step in cholinergic  
 CC neurotransmitter biosynthesis and regeneration (cholinergic transmissions  
 CC are crucial to brain functions such as learning and memory). The present  
 CC sequence encodes human HACT  
 XX  
 SQ Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;

Query Match 99.7%; Score 1738.2; DB 8; Length 1743;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60  
 Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120

Db	61	 GTTGGAATATGGGCTGCCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Db	361	 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTACGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	 CTCTATTCTGTGGCCTACACTGATGTCGTTACGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG	660
Db	601	 ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	 CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTAGAGG	780
Db	721	 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTTCGGG	840
Db	781	 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960

Db 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 |||

Db 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080  
 |||

Db 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 |||

Db 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 |||

Db 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 |||

Db 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 |||

Db 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 |||

Db 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||

Db 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440

Qy 1441 AACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||

Db 1441 AACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||

Db 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||

Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTTACCAATAAA 1680  
 |||

Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTTACCAATAAA 1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||

Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA 1740

Qy 1741 TGA 1743  
 |||

Db 1741 TGA 1743

RESULT 6

AAF81711

ID AAF81711 standard; cDNA; 1743 BP.

XX

AC AAF81711;

XX

DT 01-JUN-2001 (first entry)

XX

DE Rat high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;  
KW ss.

XX

OS Rattus norvegicus.

XX

FH Key Location/Qualifiers

FT CDS 1. .1743

FT /\*tag= a

FT /product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

PR 27-DEC-1999; 99JP-00368991.

XX

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

PI Haga T, Okuda T;

XX

DR WPI; 2001-226688/23.

DR P-PSDB; AAB74664.

XX

PT New rat and human spinal cord high affinity choline transporters, useful  
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for  
PT treating Alzheimer's disease.

XX

PS Claim 6; Page 64-68; 90pp; Japanese.

XX

CC The present sequence encodes a rat (Rattus norvegicus) high affinity  
CC choline transporter protein designated cho-1. The cho-1 protein has  
CC nootropic and neuroprotective activities. The cho-1 polynucleotide and  
CC protein can be used for the diagnosis of diseases related to the  
CC expression of cho-1 by comparing the cho-1 polynucleotide sequence in a  
CC sample to that of a control. Drug compositions containing the cho-1  
CC protein or expression promoters or inhibitors of cho-1 are useful for  
CC treating disorders characterised by abnormal levels of cho-1, such as  
CC Alzheimer's disease

XX

SQ Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;

Query Match

80.0%; Score 1394.2; DB 4; Length 1743;

Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| ||||| ||||| ||| || ||| ||||| ||||| || || |||
Db      1 ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTCTG 60

Qy     61 GTTGGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 GTTGGAATATGGGCTGCATGGAAAACCAAAACAGCGGTAATGCAGAAGAACGCAGCGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTGGTTTACCATGACAGCC 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||
Db    181 ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    301 TTTT'TTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTTC 420
      ||||| ||||| || || || ||||| || ||||| ||||| ||||| |||||
Db    361 ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| ||||| || || || ||||| ||||| ||||| |||||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || |||| |||| ||||| ||||| || || || || |||||
Db    481 GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG 600

Qy    601 ATCAGCGTCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAC'TGCTGTG 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    601 ATCAGTGTCCATT'TGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG 660

Qy    661 CATGCCAAATACCAAAAGCCGTGGCTGGGAAC'TGTTGACTCATCTGAAGTCTACTCTTGG 720
      ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
Db    661 CATGCTAAATACAGAGTCCCTGGCTGGGAACCAT'TGAATCAGTTGAAGTCTACACCTGG 720

Qy    721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    721 CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG 780

Qy    781 GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG 840
      || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	781	GTCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCAGCCATTTCGATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTTCTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTCATGTCCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTCTACATCATCATCTTCCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTTTTCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCCAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT	1620
Qy	1621	GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAACTTGCACCTGTAAAGCCTCGACAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAA	1680

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| ||||||||||||||||||||||||| || ||||||||||||||||||| |||||  
 Db 1681 GAGGCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA 1740

Qy 1741 TGA 1743  
 |||  
 Db 1741 TGA 1743

RESULT 7  
 ADD50642

ID ADD50642 standard; cDNA; 4904 BP.

XX

AC ADD50642;

XX

DT 15-JAN-2004 (first entry)

XX

DE cDNA encoding rat high-affinity choline transporter (rCHT).

XX

KW Rat; high-affinity choline transporter; rCHT; cholinergic function;

KW Parkinson's disease; Huntington's disease; Alzheimer's disease;

KW schizophrenia; dysautonomia; myasthenia gravis; brain;

KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;

KW neuroprotective; neuroleptic; gene; ss.

XX

OS Rattus sp.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	224. .1966
----	-----	------------

FT		/*tag= a
----	--	----------

FT		/product= "rCHT"
----	--	------------------

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX

PF 23-JUL-2001; 2001US-00911077.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (BLAK/) BLAKELY R D.

PA (APPA/) APPARSUNDARAM S.

PA (FERG/) FERGUSON S.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX

DR WPI; 2003-810914/76.

DR P-PSDB; ADD50643.

XX

PT Novel isolated polynucleotide encoding human or mouse high affinity

PT choline transporter polypeptide, useful in gene therapy to increase

PT cholinergic function in a cell of a patient suffering from Alzheimer's

PT disease.

XX

PS Example 1; SEQ ID NO 5; 74pp; English.

XX





Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 ||||| ||||| || ||||| || ||||| || ||||| ||||| |||||  
 Db 764 CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG 823

Qy 601 ATCAGCGTCCCCCTTTCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 824 ATCAGTGTCCCATTTGCCCTGTACATCCTGCAGTCACCGACATTGGATTCACTGTGTG 883

Qy 661 CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 884 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 943

Qy 721 CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 944 CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG 1003

Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTTGGCAGCTTTCGGG 840  
 || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1004 GTCCTCTCTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCCTTGGCAGCTTTTGGG 1063

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1064 TGCCTGGTGATGGCTCTACCAGCCATTTGCATTGGGGCCATTGGAGCCTCCACAGACTGG 1123

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1124 AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC 1183

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGAGTT 1020  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1184 CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGCTGTT 1243

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA 1080  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1244 TCTGCTGCTGTCATGTCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT 1303

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1304 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC 1363

Qy 1141 ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1364 ATGAGGATCACTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG 1423

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1424 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG 1483

Qy 1261 CTGCTTTGTGTAATCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1484 CTGCTCTGTGTAATCTTCAACGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1543

Qy 1321 TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1544 TTTGGACTTTTCTGAGAATTACCGGAGGAGAGCCATATCTATACTTGACGCCCTTAATC 1603

Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1604	TTCTACCCTGGTTATTACCCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1663
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1664	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCCTATCTAGCCAAGTAT	1723
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1724	CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATATATTTGATGCTGTTGTCTCAAGG	1783
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1784	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAACATCAAATTAAAT	1843
Qy	1621	GAACTTGCACCTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTACCAATAAA	1680
Db	1844	GAACTTGCACCTGTAAAGCCTCGACAGAGCCTAACCTCAGTTCAACTTTACCAATAAA	1903
Qy	1681	GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1904	GAGGCCTCTCCTTGATGTTGATTCCAGTCCAGAGGGATCTGGGACTGAAGATAACTTACAA	1963
Qy	1741	TGA	1743
Db	1964	TGA	1966

ID ADD50640 standard; cDNA; 1743 BP.

AC ADD50640;

DT 15-JAN-2004 (first entry)

DE cDNA encoding mouse high-affinity choline transporter (mCHT) #1.

KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;  
KW Parkinson's disease; Huntington's disease; Alzheimer's disease;  
KW schizophrenia; dysautonomia; myasthenia gravis; brain;  
KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
KW neuroprotective; neuroleptic; gene; ss.

OS Mus sp.

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1.1743
----	-----	--------

FT /\*tag= a

FT /product= "mCHT #1"

PN US2003114399-A1.

PD 19-JUN-2003.

PF 23-JUL-2001; 2001US-00911077.

```

XX PR      23-JUL-2001; 2001US-00911077.
XX
PA      (BLAK/) BLAKELY R D.
PA      (APPA/) APPARSUNDARAM S.
PA      (FERG/) FERGUSON S.
XX
PI      Blakely RD, Apparsundaram S, Ferguson S;
XX
DR      WPI; 2003-810914/76.
DR      P-PSDB; ADD50641.
XX
PT      Novel isolated polynucleotide encoding human or mouse high affinity
PT      choline transporter polypeptide, useful in gene therapy to increase
PT      cholinergic function in a cell of a patient suffering from Alzheimer's
PT      disease.
XX
PS      Claim 30; SEQ ID NO 3; 74pp; English.
XX
CC      The present invention relates to the isolation of polynucleotide
CC      sequences encoding human and mouse high-affinity choline transporter
CC      (hCHT and mCHT respectively), and the proteins they encode. The gene
CC      encoding hCHT is located on chromosome 2ql2. The polynucleotide sequence
CC      encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
CC      polynucleotide sequence when delivered to a cell, increases cholinergic
CC      function in the cell that is in a patient having Parkinson's disease,
CC      Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC      myasthenia gravis. The hCHT antibody is useful for controlling
CC      transporter CHT proteins to the brain, and for treating the above
CC      mentioned diseases. The antibody is also useful for diagnosing the above
CC      mentioned disorders and to detect the influence of cholinergic
CC      signalling. The present sequence encodes mCHT. Note: The sequence data
CC      for this patent was obtained in electronic format directly from the USPTO
CC      web site at seqdata.uspto.gov.
XX
SQ      Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match          78.9%;   Score 1375;   DB 9;   Length 1743;
Best Local Similarity 86.8%;   Pred. No. 0;
Matches 1513; Conservative    0; Mismatches 230; Indels    0; Gaps    0;

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60
Db              ||| |||||||||||||||| |||||| | || | |||||| || || || |||
Qy      61 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
Db              |||||||||||||||| ||| |||||||||||| ||| | |||||||||||| |||
Qy      121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
Db              ||||||||| || |||| || |||||| |||||||| |||||||||||||
Qy      181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
Db              |||||||| |||||| || |||||| ||||| |||||| |||| |||||| |
Qy      181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 240

```

Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTT	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCTGTCTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTATCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTTCTTCAGACAAAATGCTTCGGACAAAAGAAATCGTTTGGGTT	1140



XX  
 KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;  
 KW Parkinson's disease; Huntington's disease; Alzheimer's disease;  
 KW schizophrenia; dysautonomia; myasthenia gravis; brain;  
 KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
 KW neuroprotective; neuroleptic; gene; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1. .1743  
 FT /\*tag= a  
 FT /product= "mCHT #2"  
 XX  
 PN US2003114399-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 23-JUL-2001; 2001US-00911077.  
 XX  
 PR 23-JUL-2001; 2001US-00911077.  
 XX  
 PA (BLAK/) BLAKELY R D.  
 PA (APPA/) APPARSUNDARAM S.  
 PA (FERG/) FERGUSON S.  
 XX  
 PI Blakely RD, Apparsundaram S, Ferguson S;  
 XX  
 DR WPI; 2003-810914/76.  
 DR P-PSDB; ADD50661.  
 XX  
 PT Novel isolated polynucleotide encoding human or mouse high affinity  
 PT choline transporter polypeptide, useful in gene therapy to increase  
 PT cholinergic function in a cell of a patient suffering from Alzheimer's  
 PT disease.  
 XX  
 PS Example 4; SEQ ID NO 23; 74pp; English.  
 XX  
 CC The present invention relates to the isolation of polynucleotide  
 CC sequences encoding human and mouse high-affinity choline transporter  
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
 CC polynucleotide sequence when delivered to a cell, increases cholinergic  
 CC function in the cell that is in a patient having Parkinson's disease,  
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
 CC myasthenia gravis. The hCHT antibody is useful for controlling  
 CC transporter CHT proteins to the brain, and for treating the above  
 CC mentioned diseases. The antibody is also useful for diagnosing the above  
 CC mentioned disorders and to detect the influence of cholinergic  
 CC signalling. The present sequence encodes mCHT. Note: The sequence data  
 CC for this patent was obtained in electronic format directly from the USPTO  
 CC web site at seqdata.uspto.gov.  
 XX  
 SQ Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match 78.9%; Score 1375; DB 9; Length 1743;

Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| ||||| ||||| ||||| ||| | ||||| || || || |||
Db      1 ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTTCTG 60

Qy     61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db     61 GTTGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| |||
Db    121 GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTTACCATGACAGCC 180

Qy    181 ACCTGGGTCCGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| |||
Db    181 ACCTGGGTGGGAGGAGGTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    241 GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    301 TTTTTCGCAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGT'TTATTCCCTGCACTGATGGGAGAAATGTTTC 420
      ||||| ||||| || || || || || || ||||| ||||| ||||| |||
Db    361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTCAATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || |||| ||||| ||||| ||||| || || || |||| |||
Db    481 GTGAACATATCGGTCAATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTT'TGTAGGGCTGTGG 600
      ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    601 ATCAGTGTCCCTTTTGGCCTGTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 660

Qy    661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGACTCATCTGAAGTCTACTCTTGG 720
      ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||
Db    661 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 720

Qy    721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db    721 CTTGATAAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 780

Qy    781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG 840
      || ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||
```

Db 781 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG 840

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 841 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 900

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 901 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC 960

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 961 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1020

Qy 1021 TCTGCTGCTGTTATGTTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1021 TCAGCTGCTGTGATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1080

Qy 1081 CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1081 CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC 1140

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 ||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1141 ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTGCTGACGAAG 1200

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1201 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCACAG 1260

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1261 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1320

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 | ||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 1321 TTTGGACTATTCTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGAGCCCTTAATC 1380

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1381 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTAAA 1440

Qy 1441 AACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1441 ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1500

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1501 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTGCAAGG 1560

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1561 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1620

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1621 GAACTTGCACTGTGAAACCTCGGCAGAGCCTAACCCTCAGTTCAACTTTCACCAATAAG 1680



Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| |||||||||||||||||||| || ||||||||||||||||||||  
 Db 1681 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAATTTACAA 1740  
 Qy 1741 TGA 1743  
 |||  
 Db 1741 TGA 1743

RESULT 10

AAF81713

ID AAF81713 standard; cDNA; 1743 BP.

XX

AC AAF81713;

XX

DT 01-JUN-2001 (first entry)

XX

DE Mouse high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;  
 KW ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 1. .1743

FT /\*tag= a

FT /product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

PR 27-DEC-1999; 99JP-00368991.

XX

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

PI Haga T, Okuda T;

XX

DR WPI; 2001-226688/23.

DR P-PSDB; AAB74666.

XX

PT New rat and human spinal cord high affinity choline transporters, useful  
 PT in diagnosis of Alzheimer's disease and screening promoters as drugs for  
 PT treating Alzheimer's disease.

XX

PS Claim 12; Page 78-82; 90pp; Japanese.

XX

CC The present sequence encodes a mouse (Mus musculus) high affinity choline  
 CC transporter protein designated cho-1. The cho-1 protein has nootropic and  
 CC neuroprotective activities. The cho-1 polynucleotide and protein can be  
 CC used for the diagnosis of diseases related to the expression of cho-1 by  
 CC comparing the cho-1 polynucleotide sequence in a sample to that of a

CC control. Drug compositions containing the cho-1 protein or expression  
CC promoters or inhibitors of cho-1 are useful for treating disorders  
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease  
XX  
SQ Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;

Query Match 78.8%; Score 1373.4; DB 4; Length 1743;  
Best Local Similarity 86.7%; Pred. No. 0;  
Matches 1512; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```
Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60
      ||| ||||| || ||||| ||||| ||| | ||||| || || || |||
Db      1 ATGTCTTTCCACGTAGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG 60

Qy     61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 GTTGAATATGGGCTGCATGGAAAACCAAAACAGCGGCAACCCAGAAGAGCACAGTGAA 120

Qy    121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180
      ||||| ||||| || ||||| || ||||| ||||| ||||| |||||
Db    121 GCCATCATAGTCGGGGCCGTGACATTGGTTTGTGGTTGGTGGTTTACCATGACAGCC 180

Qy    181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db    181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT 240

Qy    241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    241 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG 300

Qy    301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    301 TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTTCACAG 360

Qy    361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTCTGCACTGATGGGAGAAATGTTT 420
      ||||| ||||| || || || || || ||||| ||||| ||||| |||||
Db    361 ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTT 420

Qy    421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480
      ||||| ||||| ||||| || || ||||| ||||| ||||| ||||| |||||
Db    421 TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT 480

Qy    481 ATGCACATTTCTGTTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540
      || |||| || |||| ||||| ||||| ||||| ||||| ||||| |||||
Db    481 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 540

Qy    541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600
      ||||| ||||| || ||||| || ||||| || ||||| ||||| |||||
Db    541 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 600

Qy    601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTCTGTG 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    601 ATCAGTGTCCCTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 660

Qy    661 CATGCCAAATACCAAAAGCCGTGGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
      ||||| ||||| || || ||||| ||||| ||||| ||||| |||||
Db    661 CATGCTAAATACCAGAGTCCCTGGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG 720
```

Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGCTTCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCCTGT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTTCGGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTCTACATCATCATCTTCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACTCTCTCCATGGTTACCTCATTCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCCTCAAATTAGATGTATTTGATGCTGTTGTCGCAAGG	1560

Qy	1561	CACAGTGAAGAAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620
Db	1561	CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT	1620
Qy	1621	GAAC TTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA	1680
Db	1621	GAAC TTGCACTTGTGAAACCTCGGCAGAGCCTAACCCCTCAGTTCAACTTTCACCAATAAG	1680
Qy	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG	1740
Db	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCGGAGGGGTCTGGGACTGAAGATAACTTACAA	1740
Qy	1741	TGA	1743
Db	1741	TGA	1743



Db 727 GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG 786  
 Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 Db 787 CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG 846  
 Qy 601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG 660  
 Db 847 ATCAGTGTCCCTTTTGCCCTGTACATCCTGCAGTCACCGACATCGGATTACAGCTGTG 906  
 Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG 720  
 Db 907 CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATTGAATCAGTTGAAGTCTACACCTGG 966  
 Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 Db 967 CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG 1026  
 Qy 781 GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTCGGG 840  
 Db 1027 GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCTTCCTGGCAGCTTTCGGG 1086  
 Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 Db 1087 TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG 1146  
 Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 Db 1147 AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATCTC 1206  
 Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 Db 1207 CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGTCTGTT 1266  
 Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCATGTTTGCA 1080  
 Db 1267 TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCTATGTTTGCT 1326  
 Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGACAAAAGAAATCGTTTGGGTT 1140  
 Db 1327 CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATTTGTGTGGGTC 1386  
 Qy 1141 ATGCGAATCACAGTGTGTTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 Db 1387 ATGAGGATCACTGTGCTTGTGTTGCGAGCATCTGCAACAGCCATGGCTTTGCTGACGAAG 1446  
 Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG 1260  
 Db 1447 ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGCTTACATCATCATCTTCCCACAG 1506  
 Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 Db 1507 CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT 1566  
 Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380

Db 1567 TTTGGACTATTCTTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC 1626

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTAAAA 1440  
 ||||| ||||| ||||| | ||||| ||||| || |||||

Db 1627 TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA 1686

Qy 1441 AACTTGGCATGGTTACATCATCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 || || ||||| ||||| ||||| ||||| ||||| |||||

Db 1687 ACTCTCTCCATGGTTACCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT 1746

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 ||||| ||||| ||||| || ||||| ||||| ||||| |||||

Db 1747 CTATTTGAAAGTGGAACCTTGCCCTCCAAAATTAGATGTATTTGATGCTGTTGTCGCAAGG 1806

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

Db 1807 CACAGTGAAGAGAACATGGACAAGACCATTCTAGTCAGAAATGAAAATATCAAATTAAAT 1866

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 ||||| ||||| || || ||||| | ||||| ||||| |||||

Db 1867 GAACTTGCACTGTGAAACCTCGGCAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG 1926

Qy 1681 GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 ||||| ||||| ||||| ||||| || ||||| ||||| |||||

Db 1927 GAGGCCCTCCTTGATGTTGATTCCAGTCCGGAGGGTCTGGGACTGAAGATAACTTACAA 1986

Qy 1741 TGA 1743  
 |||

Db 1987 TGA 1989

# RESULT 12

AAH49206

ID AAH49206 standard; DNA; 8760 BP.

XX

AC AAH49206;

XX

DT 26-NOV-2001 (first entry)

XX

DE Human CHOT exons 6, 7, 8 and 3' UTR region DNA.

XX

KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;

KW neuroprotective; gene therapy; antisense therapy; degenerative disease;

KW cognitive disorder; Alzheimer's disease; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .4853

FT /\*tag= a

FT /product= "CHOT"

FT /note= "This sequence is interrupted by introns"

FT exon 41. .194

FT /\*tag= b

FT /number= 6

FT intron 195. .2456

FT /\*tag= c

FT /number= 6  
 FT exon 2455. .2674  
 FT /\*tag= d  
 FT /number= 7  
 FT intron 2675. .4223  
 FT /\*tag= e  
 FT /number= 7  
 FT exon 4224. .4853  
 FT /\*tag= f  
 FT /number= 8  
 FT 3'UTR 4854. .8760  
 FT /\*tag= g

XX

PN DE10009055-A1.

XX

PD 30-AUG-2001.

XX

PF 28-FEB-2000; 2000DE-01009055.

XX

PR 28-FEB-2000; 2000DE-01009055.

XX

PA (BRUE/) BRUESS M.

PA (BOEN/) BOENISCH H.

XX

PI Bruess M, Boenisch H;

XX

DR WPI; 2001-590709/67.

XX

PT A new gene encoding human choline transporter, designated hCHOT is  
 PT located on chromosome 2q11-13 and is useful to treat degenerative  
 PT disorders such as Alzheimer's disease.

XX

PS Disclosure; Page 9-11; 12pp; German.

XX

CC This invention describes a novel gene encoding human choline transporter,  
 CC designated hCHOT which is located on chromosome 2q11-13. The products of  
 CC the invention have nootropic and neuroprotective activity and can be used  
 CC for gene or antisense therapy. (I) is used to treat degenerative disease,  
 CC particularly cognitive disorders such as Alzheimer's disease. Sense and  
 CC antisense oligonucleotides derived from the gene may be used in  
 CC diagnostics and other techniques. This sequence represents exons 6-8 and  
 CC the 3' UTR fragment encoding the human CHOT protein described in the  
 CC invention

XX

SQ Sequence 8760 BP; 2727 A; 1619 C; 1565 G; 2849 T; 0 U; 0 Other;

Query Match 36.2%; Score 630.8; DB 5; Length 8760;  
 Best Local Similarity 99.7%; Pred. No. 1.8e-172;  
 Matches 632; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGC 1169  
 | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 4220 ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGC 4279  
 Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAAACGTGTATGGGCTCTGGTACCTCAGTTC 1229  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 4280 ATCTGCAACAGCCATGGCCTTGCTGACGAAAACGTGTATGGGCTCTGGTACCTCAGTTC 4339



Qy	1230	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC	1289
Db	4340	TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC	4399
Qy	1290	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	1349
Db	4400	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	4459
Qy	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	1409
Db	4460	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	4519
Qy	1410	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTTCTTAAC	1469
Db	4520	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTTCTTAAC	4579
Qy	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	4580	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	4639
Qy	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	4640	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	4699
Qy	1590	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	4700	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	4759
Qy	1650	CATGACCCTCAGCTCAACTTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	1709
Db	4760	CATGACCCTCAGCTCAACTTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	4819
Qy	1710	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	1743
Db	4820	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	4853

# RESULT 13

ADD50656

ID ADD50656 standard; DNA; 119040 BP.

XX

AC ADD50656;

XX

DT 15-JAN-2004 (first entry)

XX

DE BAC sequence containing hCHT gene.

XX

KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;

KW cholinergic function; Parkinson's disease; Huntington's disease;

KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;

KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;

KW nootropic; neuroprotective; neuroleptic; bacterial artificial chromosome;

KW BAC; ds.

XX

OS Homo sapiens.

XX

PN US2003114399-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 23-JUL-2001; 2001US-00911077.  
 XX  
 PR 23-JUL-2001; 2001US-00911077.  
 XX  
 PA (BLAK/) BLAKELY R D.  
 PA (APPA/) APPARSUNDARAM S.  
 PA (FERG/) FERGUSON S.  
 XX  
 PI Blakely RD, Apparsundaram S, Ferguson S;  
 XX  
 DR WPI; 2003-810914/76.  
 XX  
 PT Novel isolated polynucleotide encoding human or mouse high affinity  
 PT choline transporter polypeptide, useful in gene therapy to increase  
 PT cholinergic function in a cell of a patient suffering from Alzheimer's  
 PT disease.  
 XX  
 PS Example 3; SEQ ID NO 19; 74pp; English.  
 XX  
 CC The present invention relates to the isolation of polynucleotide  
 CC sequences encoding human and mouse high-affinity choline transporter  
 CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
 CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
 CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
 CC polynucleotide sequence when delivered to a cell, increases cholinergic  
 CC function in the cell that is in a patient having Parkinson's disease,  
 CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
 CC myasthenia gravis. The hCHT antibody is useful for controlling  
 CC transporter CHT proteins to the brain, and for treating the above  
 CC mentioned diseases. The antibody is also useful for diagnosing the above  
 CC mentioned disorders and to detect the influence of cholinergic  
 CC signalling. The present sequence represents a bacterial artificial  
 CC chromosome (BAC) sequence containing the hCHT gene. Note: The sequence  
 CC data for this patent was obtained in electronic format directly from the  
 CC USPTO web site at seqdata.uspto.gov.  
 XX  
 SQ Sequence 119040 BP; 37072 A; 22876 C; 21708 G; 36882 T; 0 U; 502 Other;

Query Match 36.2%; Score 630.8; DB 9; Length 119040;  
 Best Local Similarity 99.7%; Pred. No. 7.1e-172;  
 Matches 632; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGC 1169  
 | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 30755 ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGC 30814  
 Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 1229  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 30815 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 30874  
 Qy 1230 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 1289  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 30875 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 30934

Qy	1290	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	1349
Db	30935	CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGG	30994
Qy	1350	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	1409
Db	30995	GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA	31054
Qy	1410	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC	1469
Db	31055	TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC	31114
Qy	1470	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	1529
Db	31115	CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA	31174
Qy	1530	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	1589
Db	31175	ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT	31234
Qy	1590	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	1649
Db	31235	TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG	31294
Qy	1650	CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	1709
Db	31295	CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC	31354
Qy	1710	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	1743
Db	31355	AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA	31388

# RESULT 14

ADD50651

ID ADD50651 standard; DNA; 142299 BP.

XX

AC ADD50651;

XX

DT 15-JAN-2004 (first entry)

XX

DE BAC sequence #2 containing hCHT DNA.

XX

KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;  
 KW cholinergic function; Parkinson's disease; Huntington's disease;  
 KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;  
 KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;  
 KW nootropic; neuroprotective; neuroleptic; bacterial artificial chromosome;  
 KW BAC; ds.

XX

OS Homo sapiens.

XX

PN US2003114399-A1.

XX

PD 19-JUN-2003.

XX



```

Qy      1350 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 1409
          |||
Db      94913 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 94972

Qy      1410 TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC 1469
          |||
Db      94973 TGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAAC 95032

Qy      1470 CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 1529
          |||
Db      95033 CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 95092

Qy      1530 ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 1589
          |||
Db      95093 ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 95152

Qy      1590 TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG 1649
          |||
Db      95153 TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG 95212

Qy      1650 CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 1709
          |||
Db      95213 CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 95272

Qy      1710 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 1743
          |||
Db      95273 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 95306

```

RESULT 15

AAF81710

ID AAF81710 standard; cDNA; 1731 BP.

XX

AC AAF81710;

XX

DT 01-JUN-2001 (first entry)

XX

DE C. elegans high affinity choline transporter protein encoding cDNA.

XX

KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;  
KW ss.

XX

OS Caenorhabditis elegans.

XX

FH Key Location/Qualifiers

FT CDS 1. .1731

FT /\*tag= a

FT /product= "high affinity choline transporter"

XX

PN WO200116315-A1.

XX

PD 08-MAR-2001.

XX

PF 18-AUG-2000; 2000WO-JP005545.

XX

PR 27-AUG-1999; 99JP-00240642.

XX

XX

XX

DR

XX

XX

XX

XX

Best Local Similarity 55.1%; Pred. No. 5.5e-95;

Matches 862; Conservative 0; Mismatches 637; Indels 66; Gaps 5;

[illegible]

Db 370 GGCCAACGAATCGGTGGCTTGATGTATGTTCCAGCACTTCTTGGTGAAACATTCTGGACA 429

Qy 427 GCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCAC 486  
 ||||| ||| | || | | || | | | | | | | | | | | | |

Db 430 GCAGCCATTCTTTCCGGCACTTGGTGCAACACTGTCGGTAATTCTTGGAAATCGACATGAAT 489

Qy 487 ATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTAT 546  
 || || | | | || | | | ||||| | ||||| | | || | ||||

Db 490 GCATCAGTGACCCTGTCGGCCTGTATTGCCGTATTCTACACATTCACCGGTGGATACTAT 549

Qy 547 TCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGC 606  
 | || || ||||| ||||| || ||||| || || ||||| | ||

Db 550 GCAGTCGCGTACACTGACGTGCTTCAACTATTTTGCATTTTCGTGCGTTTGTGGGTTTGC 609

Qy 607 GTCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTGCATGCC 666  
 || || | | || | || | | || | | | || | |

Db 610 GTGCCGCGGCTATGGTGCATGATGGTGCAGGATATTTCCAGGAATGCAG----- 661

Qy 667 AAATACCAAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGGCTTGAT 726  
 | | | | | | | | | | | | | | | | | | | | |

Db 662 -----GCGACTGGATTGGAGAGATTGGAGGATTCAAAGAAACATCTCTCTGGATTGAT 714

Qy 727 AGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTC 786  
 | | || | | | | || ||||| ||||| || || |||||

Db 715 TGCATGCTTCTCCTTGTCTTTGGAGGAATTCATGGCAAGTGTACTTCCAAAGAGTTCTC 774

Qy 787 TCTTCTTCCTCAGCCACCTATGCTCAAGTGTGCTTCTCCTGGCAGCTTTCGGGTGCTG 846  
 || || | || | || || | |||| | || | | |||| || |

Db 775 TCCTCAAAAAGTCTCATGGAGCACAGACGTTGTGCTTGTGGCGGGCGTCGGATGCATT 834

Qy 847 GTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAG 906  
 | ||||| || ||| | | || || || || || || ||||| |

Db 835 CTCATGGCGATTCCACCAGCGTTGATCGGTGCAATTGCCAGGAACACAGACTGGAGAATG 894

Qy 907 ACTGCATATGGGCTTCC-----AGATCCCAAGACTACAGAAGAGGCA 948  
 |||| ||| | | | | | | | | | | | | |

Db 895 ACTGATTATTCCCATGGAACAATGGAAGTGAAGTTCGAATCGATTCCACCGGATAAGAGA 954

Qy 949 GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT 1008  
 ||||| | | || | || | ||||| || | | |||||

Db 955 AACATGGTGGTCCCGTTGGTATTCCAGTATCTTACGCCAAGATGGGTGCGCTTTATTGGA 1014

Qy 1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGT 1068  
 || || |||| | | ||||| ||||| ||||| || | | |||||

Db 1015 CTCGGCGCAGTGTGCGCTGCTGTAATGTCATCTGCAGATTCTGTACTATCAGCAGCA 1074

Qy 1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAAGAA 1128  
 || ||||| || | ||||| |||| | | | | | || || || |||||

Db 1075 TCAATGTTTGTCTACAACATCTGGAAGCTACAATTGCCCTCACGCGTCTGAAAAAGAA 1134

Qy 1129 ATCGTTTGGGTTATGCGAATCACAGTGTGTTGTTGGAGCATCTGCAACAGCCATGGCC 1188  
 | | || ||| |||| | | | |||| |||| |||||

Db 1135 GTGATAATTGTGATGAGAATAGCCATCATCTGTGTTGGTATCATGGCAACCATCATGGCA 1194

Qy 1189 TTGCTGACGAAAAGTGTGATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTACATCGTT 1248  
 | | | | ||||| |||| | | | | | | | | | |

Db 1195 CTTACCATTCATCCATCTATGGGCTTTGGTATCTTTGTGCAGATTGGTCTACGTCATA 1254

Qy 1249 ATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTG 1308  
 ||||| || || | |||| | | | | | |||| | ||  
 Db 1255 CTCTTCCCTCAACTATTATGTGTTGTATATATGCCACGTAGCAATACGTATGGCTCATTG 1314

Qy 1309 GCAGGTTATGTTTCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTT 1368  
 || || |||| || || | | | | |||| |||| | | | ||  
 Db 1315 GCTGGCTATGCAGTCGGTCTTGTGCTCCGTTTGATTGGAGGCGAGCCACTTGTATCGCTG 1374

Qy 1369 CAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAA 1428  
 | | | | || || || || || || || | || | || |  
 Db 1375 CCAGCGTTCTTCCATTATCCAATGTATACGGATGGGG-----TACAGTAT 1419

Qy 1429 TTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTAT 1488  
 || |||| | || || || || | | || | || | || |  
 Db 1420 TTCCCATTACAGGACAACTGCTATGTTATCTTCAATGGCTACTATCTACATTGTATCAATA 1479

Qy 1489 CTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCT 1548  
 | | | | || | || || || || || || | | || | || |  
 Db 1480 CAATCGGAGAAGCTGTTCAAATCGGGACGTTTGTCTCCGGAGTGGGACGTAATGGGTTGT 1539

Qy 1549 GTTGT 1553  
 || ||  
 Db 1540 GTAGT 1544

Search completed: March 22, 2004, 12:00:55  
 Job time : 755 secs



OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 11:12:56 ; Search time 144 Seconds  
 (without alignments)  
 6717.218 Million cell updates/sec

Title: US-10-069-541-5  
 Perfect score: 1743  
 Sequence: 1 atggcctttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	1738.2	99.7	1743	4	US-09-657-252-1	Sequence 1, Appli
2	47.6	2.7	7218	1	US-08-232-463-14	Sequence 14, Appl
3	41	2.4	1857	4	US-09-640-198D-3	Sequence 3, Appli
4	41	2.4	2839	4	US-08-595-553A-1	Sequence 1, Appli
5	39.6	2.3	474	4	US-09-621-976-18033	Sequence 18033, A
6	39.2	2.2	1506	4	US-09-328-352-2245	Sequence 2245, Ap
7	39	2.2	558	4	US-09-328-352-3451	Sequence 3451, Ap
8	38.2	2.2	2028	4	US-10-162-012-28	Sequence 28, Appl
9	38.2	2.2	2326	4	US-10-162-012-26	Sequence 26, Appl
c 10	38.2	2.2	1830121	4	US-09-557-884-1	Sequence 1, Appli
c 11	38.2	2.2	1830121	4	US-09-643-990A-1	Sequence 1, Appli

	12	38	2.2	1932	4	US-09-640-198D-1	Sequence 1, Appli
	13	37.4	2.1	4344	4	US-09-601-198-165	Sequence 165, App
c	14	36.6	2.1	4160	4	US-09-134-218-1	Sequence 1, Appli
	15	36.6	2.1	148567	4	US-09-801-876B-3	Sequence 3, Appli
	16	36.6	2.1	148567	4	US-10-254-869-3	Sequence 3, Appli
	17	35.6	2.0	447	4	US-09-621-976-12063	Sequence 12063, A
c	18	35.6	2.0	2397	4	US-09-221-017B-272	Sequence 272, App
	19	35.2	2.0	1461	4	US-09-543-681A-2066	Sequence 2066, Ap
	20	35.2	2.0	2238	1	US-07-841-651-1	Sequence 1, Appli
	21	34.8	2.0	902	4	US-09-671-317-37	Sequence 37, Appl
	22	34.8	2.0	1593	4	US-09-134-001C-1673	Sequence 1673, Ap
c	23	34.8	2.0	12482	4	US-09-512-563C-25	Sequence 25, Appl
c	24	34.8	2.0	25002	4	US-08-961-527-48	Sequence 48, Appl
	25	34.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	26	34.6	2.0	561	4	US-09-107-532A-3215	Sequence 3215, Ap
	27	34.6	2.0	1005	4	US-09-107-532A-3570	Sequence 3570, Ap
	28	34.6	2.0	2847	4	US-09-484-970B-22	Sequence 22, Appl
c	29	34.6	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	30	34.2	2.0	1515	4	US-09-071-035-431	Sequence 431, App
	31	34.2	2.0	1803	4	US-09-071-035-429	Sequence 429, App
c	32	34	2.0	1109	4	US-08-956-171E-222	Sequence 222, App
c	33	34	2.0	392000	4	US-10-027-983-11	Sequence 11, Appl
c	34	33.8	1.9	369	4	US-09-543-681A-628	Sequence 628, App
c	35	33.8	1.9	3172	1	US-07-741-940-3	Sequence 3, Appli
c	36	33.8	1.9	3172	1	US-08-289-548A-3	Sequence 3, Appli
c	37	33.8	1.9	3172	1	US-08-452-654-3	Sequence 3, Appli
c	38	33.8	1.9	3172	1	US-08-452-655B-3	Sequence 3, Appli
c	39	33.8	1.9	3172	3	US-08-450-582-3	Sequence 3, Appli
c	40	33.8	1.9	3172	4	US-08-449-731-3	Sequence 3, Appli
	41	33.8	1.9	176373	3	US-09-128-155-17	Sequence 17, Appl
	42	33.6	1.9	84495	4	US-09-797-906-3	Sequence 3, Appli
	43	33.4	1.9	1626	4	US-09-328-352-602	Sequence 602, App
	44	33.4	1.9	3593	4	US-09-404-627-3	Sequence 3, Appli
	45	33.4	1.9	4205	4	US-09-404-627-1	Sequence 1, Appli

# ALIGNMENTS

```

RESULT 1
US-09-657-252-1
; Sequence 1, Application US/09657252
; Patent No. 6500643
; GENERAL INFORMATION:
; APPLICANT: Wu, Dong-Hai
; APPLICANT: Gu, Yunrong
; APPLICANT: Millard, William
; APPLICANT: He, Yun-Je
; TITLE OF INVENTION: Human High Affinity Choline Transporter cDNA
; FILE REFERENCE: MBHB00-639
; CURRENT APPLICATION NUMBER: US/09/657,252
; CURRENT FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA

```

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1743)  
US-09-657-252-1

Query Match 99.7%; Score 1738.2; DB 4; Length 1743;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1740; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT	180
Qy	181	ACCTGGGTCCGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCCGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720

Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCCTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCA	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCA	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441	ACACTTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560
Db	1501	CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA	1560

QY 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||  
 Db 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620

QY 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 |||  
 Db 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680

QY 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||  
 Db 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAA 1740

QY 1741 TGA 1743  
 |||  
 Db 1741 TGA 1743

RESULT 2

US-08-232-463-14

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 IMMU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 2.7%; Score 47.6; DB 1; Length 7218;  
Best Local Similarity 6.0%; Pred. No. 0.00054;  
Matches 23; Conservative 200; Mismatches 159; Indels 0; Gaps 0;

```
Qy      692 CTGTTGACTCATCTGAAGTCTACTCTTGGCTTGATAGTTTCTGTTGTTGATGCTGGGTG 751
      :: :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      1088 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1147

Qy      752 GAATCCCATGGCAAGCATACTTTTCTAGAGGGTTCTCTCTTCTCCTCAGCCACCTATGCTC 811
      :::: : : : : :::: :::: :::: :::: :::: :::: :::: ::::
Db      1148 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1207

Qy      812 AAGTGCTGTCCTTCTGCGAGCTTTCGGGTGCCTGGTGATGGCCATCCCAGCCATACTCA 871
      : :: :::: :::: : :::: : :::: : : : : :::: :::: ::::
Db      1208 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1267

Qy      872 TTGGGGCCATTGGAGCATCAACAGACTGGAACCAGACTGCATATGGGCTTCCAGATCCCA 931
      :: :::: : : : : : : : : : : : : : : :::: ::::
Db      1268 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1327

Qy      932 AGACTACAGAAGAGGCAGACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGT 991
      :: : : : : : : : : : : : : : : : : : :
Db      1328 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1387

Qy      992 ATATTTCTTTCTTTGGTCTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTT 1051
      : :::: :::: : : : : : : : : : : : : | || |||||
Db      1388 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAAATCTT 1447

Qy      1052 CCATCTTGTCAGCAAGTTCCAT 1073
      | ||| | | | | || |||
Db      1448 CTATCTCTTTAACTACTTGCAT 1469
```

RESULT 3

US-09-640-198D-3

; Sequence 3, Application US/09640198D  
; Patent No. 6586411  
; GENERAL INFORMATION:  
; APPLICANT: Russell, Stephen  
; APPLICANT: Kay Whye, Peng  
; TITLE OF INVENTION: System for Monitoring the Location of  
; TITLE OF INVENTION: Transgenes  
; FILE REFERENCE: 07039-295001  
; CURRENT APPLICATION NUMBER: US/09/640,198D  
; CURRENT FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 60/149,168  
; PRIOR FILING DATE: 1999-08-17  
; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1857  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-09-640-198D-3

Query Match 2.4%; Score 41; DB 4; Length 1857;  
Best Local Similarity 49.8%; Pred. No. 0.026;  
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy      404 TGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCG 463
          || |||      ||| | |   | |       | ||| ||   |   || | | | |
Db      413 TGGTGGCCACGATGCTGTATACAGGCATCGTGATCTACGCGCCTGCGCTCATCCTGAACC 472

Qy      464 TGATCATCGATGTGGATATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGT 523
          | | ||   |||| ||       || || |   | | | | ||||   ||| |||
Db      473 AAGTGACCGGGTTGGACATCTGGGCATCGCTCCTGTCCACAGGAATCATCTGCACCTTGT 532

Qy      524 ACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCA 583
          ||||   ||||| ||   | | | |||| ||   || |||||   | ||| |   |
Db      533 ACACTACCGTGGGTGGTATGAAGGCCGTGGTCTGGACAGATGTGTTCCAGGTGTGGTAA 592

Qy      584 TTTTGTAGGGCTGTGGATCAGCGTCCCC 612
          | | || ||   | ||| | | | | ||
Db      593 TGCTCGTTGGCTTCTGGGTGATCCTGGCC 621
```

#### RESULT 4

US-08-595-553A-1

; Sequence 1, Application US/08595553A  
; Patent No. 6391579  
; GENERAL INFORMATION:  
; APPLICANT: NANCY CARRASCO, ET AL.  
; TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SAME  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN  
; STREET: 90 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE  
; MEDIUM TYPE: DISKETTE  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/595,553A  
; FILING DATE: FEBRUARY 1, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG J. ARNOLD  
; REGISTRATION NUMBER: 34,287  
; REFERENCE/DOCKET NUMBER: 96700/393

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2839
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: OLIGONUCLEOTIDE
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: RAT
; INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
US-08-595-553A-1

```

```

Query Match          2.4%; Score 41; DB 4; Length 2839;
Best Local Similarity 49.8%; Pred. No. 0.035;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```

```

Qy      404 TGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCG 463
        || |||   ||| | | | | | | | | | | | | | | | |
Db      522 TGGTGGCCACGATGCTGTATACAGGCATCGTGATCTACGCGCCTGCGCTCATCCTGAACC 581

Qy      464 TGATCATCGATGTGGATATGCACATTTCTGTGCATCATCTCTGCACTCATTGCCACTCTGT 523
        | | ||   ||| | | | | | | | | | | | | | | | |
Db      582 AAGTGACCGGGTTGGACATCTGGGCATCGCTCCTGTCCACAGGAATCATCTGCACCTTGT 641

Qy      524 ACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTGCTTCAGCTCTTTTGCA 583
        ||||   |||| | | | | | | | | | | | | | | | |
Db      642 ACACTACCGTGGGTGGTATGAAGGCCGTGGTCTGGACAGATGTGTTCCAGGTTGTGGTAA 701

Qy      584 TTTTGTAGGGCTGTGGATCAGCGTCCCC 612
        | | || | | | | | | | | | | |
Db      702 TGCTCGTTGGCTTCTGGGTGATCCTGGCC 730

```

# RESULT 5

US-09-621-976-18033

; Sequence 18033, Application US/09621976

; Patent No. 6639063

## ; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 18033

; LENGTH: 474

; TYPE: DNA



; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 16  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-18033

Query Match 2.3%; Score 39.6; DB 4; Length 474;  
Best Local Similarity 13.4%; Pred. No. 0.03;  
Matches 42; Conservative 134; Mismatches 138; Indels 0; Gaps 0;

Qy 981 CTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGAGTTTCTGCTGCTGTTATGTCATC 1040  
:: |: ::| | ||: :: | ||: | : : : ::  
Db 54 SKYCSGSYKKT TTTTWWWWTTT TKGKWARRRMSGGGKTTYMMCSKKKTKSCMAGRWKG 113  
  
Qy 1041 AGCAGATTCTTCCATCTTGTGTCAGCAAGTCCATGTTTGCACGGAACATCTACCAGCTTTC 1100  
: ::::: : |:: : : : : : : |: :: : : : : : : |:  
Db 114 KYYYSRWYYYCYKGACYMWKRWYCSSCCMMYTKGGGSMWTTTWMRRRKSYKRWTKGK 173  
  
Qy 1101 CTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTTATGCGAATCACAGTGTGTTGT 1160  
:: ::||: | : : ::::: || :|:: | | :| | : : :  
Db 174 KKKKTTWMMMAAMCYTTWRSYWMRRRAAAKTYYYCMMSKTMCCMACCCMMCMRRARS 233  
  
Qy 1161 GTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTA 1220  
: :: |: | : : : |::: | : : : : : : : :|: | : : :  
Db 234 CCMRSCMRSYTYMMCYYYMYKGRMYWWWRGGMWKRMYYKKKSMWKGSCMWKRAWW 293  
  
Qy 1221 CCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAGCTGCTTTGTGTACTCTTTGT 1280  
: :| : |: : | ::::| :| : : : : : :| : : : : : :  
Db 294 ARKTTYTWAWYYTTYKRMCCYYMRKTTYCMMWYSRWWRGSMWTARGAWWMCYWWYYY 353  
  
Qy 1281 TAAGGGAACCAACA 1294  
||::: : :| |  
Db 354 MAARKKKYMWAAAA 367

#### RESULT 6

US-09-328-352-2245

; Sequence 2245, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 2245

; LENGTH: 1506

; TYPE: DNA

; ORGANISM: Acinetobacter baumannii

US-09-328-352-2245

Query Match 2.2%; Score 39.2; DB 4; Length 1506;  
Best Local Similarity 51.1%; Pred. No. 0.085;

Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy      1008 TCTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGTCAGCAAG 1067
          ||| | ||| ||| ||||| ||||| || | | | ||| |
Db      1005 TCTAGCTGCTATTTTAGCTGCGGTTATGAGTACATTAAGCTGTCAGCTTTTGGTATGTTTC 1064

Qy      1068 TTCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGA 1127
          | | | | ||| | | | | | ||||| || | |||||
Db      1065 AAGTGCCTAACTGAAGATTTGTACAAAGGCTTCATTCGTAAAAATGCATCTCAAAAAGA 1124

Qy      1128 AATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGC 1187
          | || ||||| ||| |||| ||| ||| ||| | |||| | |||
Db      1125 GCTTGTATGGGTTGGGCGTATCATGGTGCTTGCAATTGCCGTTCTAGCAATTGTGCTTGC 1184
```

#### RESULT 7

US-09-328-352-3451

; Sequence 3451, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 3451

; LENGTH: 558

; TYPE: DNA

; ORGANISM: Acinetobacter baumannii

US-09-328-352-3451

Query Match 2.2%; Score 39; DB 4; Length 558;

Best Local Similarity 50.3%; Pred. No. 0.051;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
Qy      1455 TACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGG 1514
          || |||| | | ||| | ||| | || | || | ||| ||| |
Db      341 TAAATCAAATGATGCAAATGCTTCATGGCTGATGTTGCAAACTTCAACTTTTCAAGATG 400

Qy      1515 AACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAA 1574
          | | | || || ||| | | | | | ||| || | || | ||
Db      401 GCCGTAGTCATCTGAATGCGGCAAAGCTCAAGGTGAAGTTTCAGAAGCAAGCAGATGGAA 460

Qy      1575 CATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGT 1634
          ||||| || || || || | ||| || || | | | ||| ||
Db      461 CATGGAATTTAAACATTTCCAAACACAGAATATTTTCAGTCGTCGGTATCGCATTGGC 520

Qy      1635 GAAGCCACGAC 1645
          ||| |||
Db      521 AAAGTGAAGCC 531
```

#### RESULT 8

US-10-162-012-28

; Sequence 28, Application US/10162012  
; Patent No. 6682597  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Gu, Wei  
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS  
; FILE REFERENCE: 10448-190001  
; CURRENT APPLICATION NUMBER: US/10/162,012  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/209,845  
; PRIOR FILING DATE: 2000-06-06  
; PRIOR APPLICATION NUMBER: US 09/875,321  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: PCT/US01/18340  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/209,257  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: US 09/875,423  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: PCT/US01/18398  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/209,238  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: US 09/875,363  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: PCT/US01/18247  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/227,068  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 09/928,530  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: PCT/US01/25475  
; PRIOR FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: US 60/226,770  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 09/934,421  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26096  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/279,281  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 10/109,029  
; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: PCT/US02/09728  
; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: US 60/290,288  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US (not assigned)  
; PRIOR FILING DATE: 2002-05-13  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 2028  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-162-012-28

Query Match 2.2%; Score 38.2; DB 4; Length 2028;  
Best Local Similarity 46.1%; Pred. No. 0.22;  
Matches 239; Conservative 0; Mismatches 273; Indels 7; Gaps 3;

```
Qy      79 TGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAGTTGGTGGC 138
      |||| | ||||| | | | | ||| ||| | | ||| |
Db     126 TGGACTATGGTCCACAGTGAAGACCAAAAGAGACACAGTGAAAGGCTACTTCCTGGCTGA 185

Qy     139 CGAGATATTGGTTTATTGGTTGGTGGATTTA-CCATGACAGCTACCTGGGTCGGAGGAGG 197
      | || |||| | || |||| | | || || || || | || || ||
Db     186 AGGGAACATGGTGTGGTGGCCAGTGGGTGCATCCTTGTTTGCCAGCAATGTTGGAAGTGG 245

Qy     198 GTATATCAATGGCACAGCTGAAGCAGT'TTATGTACCAGGTATGGCCTAGCTTGGGCTCA 257
      || ||| |||| || | || | || | || | || | || | || |
Db     246 ACATTTTCATTGGCCTGGCAGGGTCAGGTGCTGCTACGGGCATTTCTGTA---TCAGCTTA 302

Qy     258 GGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAAACCTAT 317
      | || | ||| | | | | |||| || | | ||| | | | ||
Db     303 TGAAGTTAATGGCTTGTTTTCTGTGCTGATGTTGGCCTGGATCTTCCTACCCATCTACAT 362

Qy     318 GCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCAT 377
      | ||| || || || || || || || || || || || ||
Db     363 TGCTGGTCAGGTCACCACGATGCCAGAATACCTACGGAAGCGCTTCGGTGGCATCAGAAT 422

Qy     378 GGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTT 437
      | |||| | | | || | || || || || || || || ||
Db     423 CCCCATCATCCTGGCTGTACTCTACCTATTTATCTACATCTTCACCAAGATCTCGGTAGA 482

Qy     438 CTCTGCTTTGGGAGCCACCATCAGCGTGATCATCG---ATGTGGATATGCACATTTCTGT 494
      | | || |||| | ||| | | || |||| || || | | |
Db     483 CATGTATGCAGGTGCCATCTTCATCCAGCAGTCTTCGCACCTGGATCTGTACCTGGCCAT 542

Qy     495 CATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGC 554
      | || | || | |||| | | || || || || || || ||
Db     543 AGTTGGGCTACTGGCCATCACTGCTGTATACACGGTTGCTGGTGGCCTGGCTGCTGTGAT 602

Qy     555 CTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGG 593
      ||||| |||| | | || | || || || || || ||
Db     603 CTACACGGATGCCCTGCAGACGCTGATCATGCTTATAGG 641
```

RESULT 9

US-10-162-012-26

; Sequence 26, Application US/10162012

; Patent No. 6682597

; GENERAL INFORMATION:

; APPLICANT: Curtis, Rory A.J.

; APPLICANT: Silos-Santiago, Inmaculada

; APPLICANT: Gu, Wei

; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS

; FILE REFERENCE: 10448-190001

; CURRENT APPLICATION NUMBER: US/10/162,012

; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: US 60/209,845

; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: US 09/875,321

```

; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 2326
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (178)...(2202)
US-10-162-012-26

```

```

Query Match          2.2%;  Score 38.2;  DB 4;  Length 2326;
Best Local Similarity 46.1%;  Pred. No. 0.24;
Matches 239;  Conservative 0;  Mismatches 273;  Indels 7;  Gaps 3;

```

```

Qy      79 TGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCATAGTTGGTGGC 138
        |||| |   |||||   | | |   ||| |||   | |   ||| |
Db      303 TGGACTATGCTCCACAGTGAAGACCAAAAGAGACACAGTGAAAGGCTACTTCCTGGCTGA 362

Qy      139 CGAGATATTGGTTTATTGGTTGGTGGATTTA-CCATGACAGCTACCTGGGTCTGGAGGAGG 197

```

```

      | | | | | | | | | | | | | | | | | | | | | |
Db    363 AGGGAACATGGTGTGGTGGCCAGTGGGTGCATCCTTGTTTGCCAGCAATGTTGGAAGTGG 422

Qy    198 GTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTTGGGCTCA 257
      | | | | | | | | | | | | | | | | | | | | | |
Db    423 ACATTTTCATTGGCCTGGCAGGGTCAGGTGCTGCTACGGGCATTTCTGTA---TCAGCTTA 479

Qy    258 GGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAAACCTAT 317
      | | | | | | | | | | | | | | | | | | | | | |
Db    480 TGAACCTAATGGCTTGTTTCTGTGCTGATGTTGGCCTGGATCTTCCTACCCATCTACAT 539

Qy    318 GCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCAT 377
      | | | | | | | | | | | | | | | | | | | | | |
Db    540 TGCTGGTCAGGTCACCACGATGCCAGAATACCTACGGAAGCGCTTCGGTGGCATCAGAAT 599

Qy    378 GGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTT 437
      | | | | | | | | | | | | | | | | | | | | | |
Db    600 CCCCATCATCCTGGCTGTACTCTACCTATTTATCTACATCTTCACCAAGATCTCGGTAGA 659

Qy    438 CTCTGCTTTGGGAGCCACCATCAGCGTGATCATCG---ATGTGGATATGCACATTTCTGT 494
      | | | | | | | | | | | | | | | | | | | | | |
Db    660 CATGTATGCAGGTGCCATCTTCATCCAGCAGTCTTCGCACCTGGATCTGTACCTGGCCAT 719

Qy    495 CATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGC 554
      | | | | | | | | | | | | | | | | | | | | | |
Db    720 AGTTGGGCTACTGGCCATCACTGCTGTATACACGGTTGCTGGTGGCCTGGCTGCTGTGAT 779

Qy    555 CTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGG 593
      | | | | | | | | | | | | | | | | | | | | | |
Db    780 CTACACGGATGCCCTGCAGACGCTGATCATGCTTATAGG 818

```

RESULT 10

US-09-557-884-1/c

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of

; the Haemophilus influenzae Rd Genome, Fragments

; Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/557,884

; FILING DATE: 25-Apr-2000

```

;          CLASSIFICATION: <Unknown>
;
;    PRIOR APPLICATION DATA:
;
;          APPLICATION NUMBER: 08/476,102
;
;          FILING DATE: JUN-5-1995
;
;    ATTORNEY/AGENT INFORMATION:
;
;          NAME: Michelle S. Marks
;
;          REGISTRATION NUMBER: 41,971
;
;          REFERENCE/DOCKET NUMBER: PB186P3
;
;    TELECOMMUNICATION INFORMATION:
;
;          TELEPHONE: 301-309-8504
;
;          TELEFAX: 301-309-8439
;
;    INFORMATION FOR SEQ ID NO: 1:
;
;          SEQUENCE CHARACTERISTICS:
;
;            LENGTH: 1830121 base pairs
;
;            TYPE: nucleic acid
;
;            STRANDEDNESS: double
;
;            TOPOLOGY: linear
;
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

```

```

Query Match          2.2%;  Score 38.2;  DB 4;  Length 1830121;
Best Local Similarity 48.7%;  Pred. No. 19;
Matches 132;  Conservative 0;  Mismatches 138;  Indels 1;  Gaps 1;

```

```

Qy      1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAAGT 1068
          ||| || ||| || || || ||| || | | | ||| | |
Db      1428785 CTTTCCGCTATTTTAGCAGCAGTAATGAGTACATTAAGTGCAGCAATTGTTAATTCCTCT
1428726

```

```

Qy      1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
          | | || | |||| | || | | | || || || || |||
Db      1428725 AGCTCAATCACAGAAGATTTCTATAAAGGTTTTATTGCGCCCTAACGCATCTGAAAAAGAG
1428666

```

```

Qy      1129 ATCGTTTGGGTTATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGC- 1187
          |||| ||| || |||| | |||| || ||| ||| ||| | |||
Db      1428665 CTCGTATGGCTTGGCAGAATTATGGTGTAGTTATTGCCGCACTTGCTATCTGGATCGCA
1428606

```

```

Qy      1188 CTTGCTGACGAAAACGTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGT 1247
          | | ||| | | | |||| | |||| || | | ||| | ||
Db      1428605 CAAGATGAAAACAGCAAAGTATTAAACTTGTAGAATTTGCTTGGGCGGGGTTGGTAGT
1428546

```

```

Qy      1248 TATCTTCCCCCAGCTGCTTTGTGTACTCTTT 1278
          || ||| | || | | |||||
Db      1428545 GCATTGGCCCTGTTGTACTTTTCTCTCTTT 1428515

```

# RESULT 11

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

; APPLICANT: Robert D. Fleischmann

; Mark D. Adams

; Owen White

```

;           Hamilton O. Smith
;           J. Craig Venter
;   TITLE OF INVENTION: The Nucleotide sequence of
;                       the Haemophilus influenzae Rd Genome, Fragments
;                       Thereof, and Uses Thereof
;   NUMBER OF SEQUENCES: 1
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Human Genome Sciences, Inc.
;       STREET: 9410 Key West Avenue
;       CITY: Rockville,
;       STATE: MD
;       COUNTRY: USA
;       ZIP: 20850
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: 3 1/2 inch diskette
;       COMPUTER: Dell Pentium
;       OPERATING SYSTEM: MS DOS v6.22
;       SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/643,990A
;       FILING DATE: 23-Aug-2000
;       CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/487,429
;       FILING DATE: 1995-06-07
;       APPLICATION NUMBER: 08/426,787
;       FILING DATE: 1995-04-21
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Kenley K. Hoover
;       REGISTRATION NUMBER: 40,302
;       REFERENCE/DOCKET NUMBER: PB186P1C1
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 301-610-5790
;       TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 1830121 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

```

Query Match          2.2%;  Score 38.2;  DB 4;  Length 1830121;
Best Local Similarity 48.7%;  Pred. No. 19;
Matches 132;  Conservative 0;  Mismatches 138;  Indels 1;  Gaps 1;

```

```

Qy      1009 CTTGGTGCAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGT 1068
          |||  ||  |||  || || || |||  ||  |  |  |  |  |  |  |
Db      1428785 CTTTCCGCTATTTTAGCAGCAGTAATGAGTACATTAAGTGCGCAATTGTTAATTCCTCT
1428726

Qy      1069 TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA 1128
          |  |  ||  |  |||  |  ||  |  |  |  |  |  |  |  |  |
Db      1428725 AGCTCAATCACAGAAGATTTCTATAAAGGTTTATTCGCCCTAACGCATCTGAAAAAGAG
1428666

```



Qy	1129	ATCGTTTGGGTTATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGC-	1187
Db	1428665	CTCGTATGGCTTGGCAGAATTATGGTGTTAGTTATTGCCGCACTTGCTATCTGGATCGCA	
	1428606		
Qy	1188	CTTGCTGACGAAAAC TGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGT TACATCGT	1247
Db	1428605	CAAGATGAAAACAGCAAAGTATTAAACCTTG TAGAATTGCTTGGGCGGGGTTTGGTAGT	
	1428546		
Qy	1248	TATCTTCCCCCAGCTGCTTTGTGTACTCTTT	1278
Db	1428545	GCATTTGGCCCTGTTGTACTTTTCTCTCTTT	1428515

```
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellson Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 4344
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-165
```

```
Query Match          2.1%; Score 37.4; DB 4; Length 4344;
Best Local Similarity 48.0%; Pred. No. 0.64;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
```

```
Qy      1401 TGATGATAATGGTATATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATC 1460
      || ||| |      || | |      || | |      || || | |      |      | | |
Db      864 TGTTCATCAATCAGTAGATTTTTTAAAAGTAAATATTGAAGCATTAAATTAATCATCAACC 923

Qy      1461 ATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTT 1520
      | |      || || |      || | |      | | |      || | | |      || | | | | | |
Db      924 ACTTAAAAACACAACATGAAACGATTTTATTAACAAAAATGTTACAGATATTAGTGCTTT 983

Qy      1521 GCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGA 1580
      || || | |      | || || | || ||      | | | | || || | || |
Db      984 AAGTAACTTATTAGAAATTTTGAAGACTAATGAAATTACAAATAATGAATGAAACCAATT 1043

Qy      1581 TAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGATGAA 1623
      | || || | | | || || || | || || |      | || ||
Db      1044 AATTACGATTTTAATTAATCATGCACCTATTGATAAAATTGAA 1086
```

#### RESULT 14

```
US-09-134-218-1/c
; Sequence 1, Application US/09134218A
; Patent No. 6312926
; GENERAL INFORMATION:
; APPLICANT: Shatkin, Aaron J.
; APPLICANT: Pillutla, Renuka
; APPLICANT: Reinberg, Danny
; APPLICANT: Yu, Zheng
; APPLICANT: Moldanado, Edio
; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
```

```
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4160
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-134-218-1
```

```
Query Match          2.1%; Score 36.6; DB 4; Length 4160;
Best Local Similarity 46.6%; Pred. No. 1.1;
Matches 117; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
```

```
Qy      1417 TATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAACCAACATT 1476
          || || || || || || || || || || || || || || || || || || || || ||
Db      2323 TAAACTTAAAATCCAACATTTAAAAAACTCAATATGCTTACAGCTTCAGATTGCTAGTTA 2264

Qy      1477 TGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGAT 1536
          || || || || || || || || || || || || || || || || || || || || ||
Db      2263 TGAATCAAATGTAAAGGTATCTATTACATACAAACAGGCTCTATTTTCATTAACCTCAAT 2204

Qy      1537 GTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAATTCTTGTC 1596
          | || | || | || | || | || | || | || | || | || | || | ||
Db      2203 CTGATTTAACCTTTGGGTATTTCAATCTGTAGACTCCACAGGGTAAGGCTGAATTTATTC 2144

Qy      1597 AAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAGCATGACC 1656
          | || || || || || | | | || || || | | | || | || | || | ||
Db      2143 AGGTATAAATAAAATATTTAGGTCCATGATGTACTGTAGTTCCAAGGAAACCAAATGTAC 2084

Qy      1657 CTCAGCTCAAC 1667
          | | | ||
Db      2083 CAAATATATAC 2073
```

# RESULT 15

```
US-09-801-876B-3
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3
```

Query Match 2.1%; Score 36.6; DB 4; Length 148567;  
 Best Local Similarity 46.1%; Pred. No. 12;  
 Matches 123; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

```

Qy      1416 ATATAATCAGAAATTTCCATTTAAAACACTTGCCATGGTTACATCATTCTTAACCAACAT 1475
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31673 AGATAATCAGTTGTTTTAACTTTTAATTTAAGCAGTAGCAGAATGACTTTTTGGGAACTT 31732

Qy      1476 TTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAAATTAGA 1535
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31733 AGGAATTTGGAAACCTTTTTATTCTATGTATTGAATATCAACTATGTAATTTAGTCTAAG 31792

Qy      1536 TGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAAACATGGATAAGACAATTCTTGT 1595
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31793 GTTATATGCTAGAAACATTTCAAAAACGAAAGCAGCAGCAATGACATCAAAAATGCATGT 31852

Qy      1596 CAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAGCATGAC 1655
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31853 CAAAAGCAAATGGTTTTAAATAGAAATACATCATTTTAACAATCTTGAAGTTTAAAAGAT 31912

Qy      1656 CCTCAGCTCAACTTTCACCAATAAAGA 1682
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31913 CCTATAAAAATCACAAACCCAGAAGGA 31939

```

Search completed: March 22, 2004, 15:19:32  
 Job time : 161 secs

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 11:50:14 ; Search time 635 Seconds  
(without alignments)  
10153.739 Million cell updates/sec

Title: US-10-069-541-5  
Perfect score: 1743  
Sequence: 1 atggcgtttccatgtggaagg.....ctgaagataatttacagtga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

%

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	1743	100.0	1743	10	US-09-911-077A-1	Sequence 1, Appli
2	1743	100.0	1813	10	US-09-911-077A-9	Sequence 9, Appli
3	1394.2	80.0	4904	10	US-09-911-077A-5	Sequence 5, Appli
4	1375	78.9	1743	10	US-09-911-077A-3	Sequence 3, Appli
5	1375	78.9	1743	10	US-09-911-077A-23	Sequence 23, Appl
6	630.8	36.2	119040	10	US-09-911-077A-19	Sequence 19, Appl
7	630.8	36.2	142299	10	US-09-911-077A-14	Sequence 14, Appl
8	376.6	21.6	1833	12	US-10-241-784-1	Sequence 1, Appli
9	363.8	20.9	1985	10	US-09-911-077A-7	Sequence 7, Appli
10	242.6	13.9	1461	9	US-09-974-300-501	Sequence 501, App
c 11	180.8	10.4	119040	10	US-09-911-077A-19	Sequence 19, Appl
c 12	180.8	10.4	142299	10	US-09-911-077A-14	Sequence 14, Appl
13	155	8.9	455	9	US-09-864-761-1838	Sequence 1838, Ap
c 14	118.6	6.8	943	15	US-10-027-632-120553	Sequence 120553,
15	72	4.1	96	9	US-09-864-761-18589	Sequence 18589, A
16	60	3.4	60	10	US-09-908-975-10249	Sequence 10249, A
17	53.8	3.1	65	10	US-09-908-975-26842	Sequence 26842, A
18	41	2.4	1857	15	US-10-428-868-3	Sequence 3, Appli
19	41	2.4	2839	9	US-09-995-007-1	Sequence 1, Appli
c 20	39.8	2.3	666	15	US-10-027-632-137101	Sequence 137101,
21	39.6	2.3	2028	9	US-09-733-630-1	Sequence 1, Appli
22	39.6	2.3	2456	9	US-09-733-630-3	Sequence 3, Appli
c 23	39.4	2.3	578	15	US-10-027-632-192644	Sequence 192644,
24	38.2	2.2	2028	9	US-09-928-530-3	Sequence 3, Appli
25	38.2	2.2	2028	14	US-10-162-012-28	Sequence 28, Appl
26	38.2	2.2	2028	15	US-10-162-102-28	Sequence 28, Appl
27	38.2	2.2	2326	9	US-09-928-530-1	Sequence 1, Appli
28	38.2	2.2	2326	14	US-10-162-012-26	Sequence 26, Appl
29	38.2	2.2	2326	15	US-10-162-102-26	Sequence 26, Appl
c 30	38.2	2.2	1830121	14	US-10-329-960-1	Sequence 1, Appli
c 31	38.2	2.2	1830121	15	US-10-329-670-1	Sequence 1, Appli
32	38	2.2	650	15	US-10-027-632-190544	Sequence 190544,
33	38	2.2	1932	15	US-10-428-868-1	Sequence 1, Appli
c 34	37.8	2.2	6306	14	US-10-239-676-129	Sequence 129, App
c 35	37.6	2.2	867	12	US-10-142-426-20	Sequence 20, Appl
c 36	37.6	2.2	867	14	US-10-123-155-20	Sequence 20, Appl
c 37	37.6	2.2	867	14	US-10-146-731-20	Sequence 20, Appl
c 38	37.6	2.2	867	14	US-10-140-472-20	Sequence 20, Appl
c 39	37.6	2.2	867	14	US-10-141-761-20	Sequence 20, Appl
c 40	37.6	2.2	867	14	US-10-142-885-20	Sequence 20, Appl
c 41	37.6	2.2	867	14	US-10-158-790-20	Sequence 20, Appl
c 42	37.6	2.2	867	15	US-10-137-871-20	Sequence 20, Appl
c 43	37.6	2.2	867	15	US-10-140-923-20	Sequence 20, Appl
c 44	37.6	2.2	867	15	US-10-141-756-20	Sequence 20, Appl
c 45	37.6	2.2	867	15	US-10-141-759-20	Sequence 20, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-911-077A-1

; Sequence 1, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

; APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/09/911,077A  
; CURRENT FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1743  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1743)  
US-09-911-077A-1

Query Match 100.0%; Score 1743; DB 10; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Db	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCCTGGAGAACCACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Db	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT	180
Qy	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTCTGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Db	361	ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCACTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	ATGCACATTTCTGTCACTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540

Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Db	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTTGTAGGGCTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Qy	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Db	1021	TCTGCTGCTGTTATGTGCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA	1080
Qy	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081	CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Qy	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141	ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Db	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Qy	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321	TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAA	1440



```

      |||
Db      1381 TTCTACCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440
Qy      1441 ACACCTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
      |||
Db      1441 ACACCTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500
Qy      1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
      |||
Db      1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560
Qy      1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
      |||
Db      1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620
Qy      1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
      |||
Db      1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680
Qy      1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
      |||
Db      1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740
Qy      1741 TGA 1743
      |||
Db      1741 TGA 1743

```

# RESULT 2

US-09-911-077A-9

; Sequence 9, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

; APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM

; APPLICANT: FERGUSON, SHAWN

; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA

; FILE REFERENCE: VBLT:008US

; CURRENT APPLICATION NUMBER: US/09/911,077A

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 1813

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (19)..(1761)

US-09-911-077A-9

Query Match 100.0%; Score 1743; DB 10; Length 1813;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG 60
      |||

```

Db 19 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 78

Qy 61 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
 |||

Db 79 GTTGAATATGGGCTGCCTGGAGAACCACAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 138

Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 180  
 |||

Db 139 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCT 198

Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
 |||

Db 199 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 258

Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
 |||

Db 259 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 318

Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
 |||

Db 319 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 378

Qy 361 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 420  
 |||

Db 379 ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTT 438

Qy 421 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 480  
 |||

Db 439 TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT 498

Qy 481 ATGCACATTTCTGTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 540  
 |||

Db 499 ATGCACATTTCTGTATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG 558

Qy 541 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 600  
 |||

Db 559 CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG 618

Qy 601 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGACATCGGGTTCACTGCTGTG 660  
 |||

Db 619 ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGACATCGGGTTCACTGCTGTG 678

Qy 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 720  
 |||

Db 679 CATGCCAAATACCAAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG 738

Qy 721 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 780  
 |||

Db 739 CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG 798

Qy 781 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGTCTCCTTCTGGCAGCTTTCGGG 840  
 |||

Db 799 GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGTCTCCTTCTGGCAGCTTTCGGG 858

Qy 841 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 900  
 |||

Db 859 TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG 918

Qy 901 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 960  
 |||  
 Db 919 AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA 978

Qy 961 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1020  
 |||  
 Db 979 CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT 1038

Qy 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1080  
 |||  
 Db 1039 TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTTCCATGTTTGCA 1098

Qy 1081 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1140  
 |||  
 Db 1099 CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT 1158

Qy 1141 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1200  
 |||  
 Db 1159 ATGCGAATCACAGTGTTTGTGTTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1218

Qy 1201 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1260  
 |||  
 Db 1219 ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG 1278

Qy 1261 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1320  
 |||  
 Db 1279 CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT 1338

Qy 1321 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1380  
 |||  
 Db 1339 TCTGGCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1398

Qy 1381 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1440  
 |||  
 Db 1399 TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA 1458

Qy 1441 ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1500  
 |||  
 Db 1459 ACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT 1518

Qy 1501 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1560  
 |||  
 Db 1519 CTATTTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTTGCAAGA 1578

Qy 1561 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1620  
 |||  
 Db 1579 CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT 1638

Qy 1621 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1680  
 |||  
 Db 1639 GAACTTGCACTTGTGAAGCCACGACAGAGCATGACCCTCAGCTCAACTTTCACCAATAAA 1698

Qy 1681 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1740  
 |||  
 Db 1699 GAGGCCTTCCTTGATGTTGATTCCAGTCCAGAAGGGTCTGGGACTGAAGATAATTTACAG 1758

Qy 1741 TGA 1743  
|||  
Db 1759 TGA 1761

RESULT 3

US-09-911-077A-5

; Sequence 5, Application US/09911077A  
; Publication No. US20030114399A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARSUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/09/911,077A  
; CURRENT FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 4904  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (224)..(1966)  
US-09-911-077A-5

Query Match 80.0%; Score 1394.2; DB 10; Length 4904;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1525; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
||| ||||||||| ||||||| |||| || ||| ||||||||| || || |||  
Db 224 ATGCCTTTCCATGTAGAAGGACTAGTAGCGATTATCCTGTTCTACCTTCTTATATTCTG 283  
  
Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
||||||| ||||||| |||| ||||||||| || | ||||||| |||||||  
Db 284 GTTGAATATGGGCTGCATGGAACCAAAAACAGCGGTAATGCAGAAGAAGCGCAGCGAA 343  
  
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 344 GCCATCATAGTTGGGGGCCGAGACATTGGTTTGTGGTTGGTGGTGGTTTACCATGACAGCC 403  
  
Qy 181 ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT 240  
||||||| ||||||| || ||||| || ||||||||| ||||||| |||||||  
Db 404 ACCTGGGTTGGAGGAGGTTACATCAACGGGACAGCTGAAGCAGTTTATGGGCCAGGTTGT 463  
  
Qy 241 GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG 300  
|| ||||||||| ||||||| ||||||||| ||||||||| ||||||||| |||||||  
Db 464 GGTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTTAGGTGGCCTG 523  
  
Qy 301 TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360  
|| ||||||||| ||||||| |||| ||||||| ||||||||| ||||||| |||  
Db 524 TTTTGTGCAAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCGTTTCAACAG 583  
  
Qy 361 ATCTATGGAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTC 420

Db	584		ATCTATGGAAAGCGCATGGGTGGGCTGCTGTTTCATCCCTGCACTGATGGGAGAGATGTTTC	643
Qy	421		TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	644		TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCTACCATCAGCGTAATCATTGATGTGGAT	703
Qy	481		ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	704		GTGAACATATCGGTCATTGTCTCCGCACTCATTGCCATTCTTTATACCCTCGTGGGAGGG	763
Qy	541		CTCTATTCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	764		CTCTACTCTGTGGCATATACTGATGTTGTACAGCTATTCTGCATTTTATAGGATTGTGG	823
Qy	601		ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	824		ATCAGTGTCCCATTTGCCCTGTCACATCCTGCAGTCACCGACATTGGATTCACTGCTGTG	883
Qy	661		CATGCCAAATACCAAAGCCGTGGCTGGGAAGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	884		CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	943
Qy	721		CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	944		CTTGATAATTTTCTGTTGTTGATGCTGGGTGGAATACCATGGCAAGCCTACTTCCAGAGG	1003
Qy	781		GTTCTCTCTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTTGGCAGCTTTCGGG	840
Db	1004		GTCTCTCTTTCATCGTCAGCGACCTATGCTCAGGTGCTGTCTTCTTGGCAGCTTTTGGG	1063
Qy	841		TGCCTGGTGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	1064		TGCCTGGTGATGGCTCTACCAGCCATTGTCATTGGGGCCATTGGAGCCTCCACAGACTGG	1123
Qy	901		AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	1124		AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGGAGGAAGCAGACATGATTCTC	1183
Qy	961		CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	1184		CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTTGGGCTTGGTGTCTGTT	1243
Qy	1021		TCTGCTGCTGTTATGTTCATCAGCAGATTCTTCCATCTTGTGCAAGTTCCATGTTTGCA	1080
Db	1244		TCTGCTGCTGTTCATGTCTCGGCTGACTCATCCATCCTATCAGCAAGTTCCATGTTTGCT	1303
Qy	1081		CGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1304		CGGAATATCTACCAGCTTTCCTTCAGACAAAATGCATCAGACAAGGAAATGTGTGGGTC	1363
Qy	1141		ATGCGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1364		ATGAGGATCACTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTCACGAAG	1423
Qy	1201		ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260



; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1743)  
US-09-911-077A-3

Query Match 78.9%; Score 1375; DB 10; Length 1743;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy	1	ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTTCTACCTTCTAATTTTGCTG	60
Db	1	ATGCCTTTCCATGTGGAAGGACTGGTAGCTATTATCCTCTTCTACCTCCTTATATTCTG	60
Qy	61	GTTGGAATATGGGCTGCCTGGAGAACCAAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA	120
Db	61	GTTGGAATATGGGCTGCATGGAAAACCAAAAACAGCGGCAACCCAGAAGAGCGCAGTGAA	120
Qy	121	GCCATCATAGTTGGTGGCCGAGATATTGGTTTATGGTTGGTGGATTACCATGACAGCT	180
Db	121	GCCATCATAGTCGGGGGCCGTGACATTGGTTTGTGGTTGGTGGTGGTTTACCATGACAGCC	180
Qy	181	ACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTAT	240
Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAACGCGATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTTC	420
Db	361	ATCTATGGAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCTCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTACATCCTGCAGTCGCAGACATCGGGTTCACTGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGCCTGTACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720

Qy	721	CTTGATAGTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721		
Db	721	CTTGATAAATTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781		
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTAAGTGCTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841		
Db	841	TGCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901		
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCACTT	1020
Db	961		
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGAGCAAGTCCATGTTTGCA	1080
Db	1021		
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTGCGCGAGTTCATGTTTGCT	1080
Qy	1081	CGGAACATCTACCAGCTTTCTTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTT	1140
Db	1081		
Db	1081	CGGAATATCTACCAGCTTTCTTTCAGACAAAATGCATCAGACAAGGAAATTGTGTGGGTC	1140
Qy	1141	ATGCGAATCAGAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA	1200
Db	1141		
Db	1141	ATGAGGATCACGTGTGCTTGTGTTGCGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAG	1200
Qy	1201	ACTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAG	1260
Db	1201		
Db	1201	ACTGTGTATGGGCTCTGGTACCTGAGCTCTGACCTTGTTTACATCATCATCTTCCCCACAG	1260
Qy	1261	CTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTT	1320
Db	1261		
Db	1261	CTGCTCTGTGTACTCTTCATCAAAGGAACCAACACTTATGGGGCAGTTGCTGGTTATATT	1320
Qy	1321	TCTGGCCTCTTCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC	1380
Db	1321		
Db	1321	TTTGGACTATTCTGAGAATTACTGGAGGAGAGCCATATCTATACTTGCAGCCCTTAATC	1380
Qy	1381	TTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAA	1440
Db	1381		
Db	1381	TTCTACCCTGGTTATTACTCTGACAAGAATGGTATATACAATCAGAGGTTCCCATTTAAA	1440
Qy	1441	ACACTTGCCATGGTTACATCATTCCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTAT	1500
Db	1441		
Db	1441	ACTCTCTCCATGGTTACCTCATTCCTTACCAACATTTGTGTTTCTTATCTAGCCAAGTAT	1500
Qy	1501	CTATTTGAAAGTGGAAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTGCAAGA	1560
Db	1501		
Db	1501	CTATTTGAAAGTGGAAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGTGCAAGG	1560
Qy	1561	CACAGTGAAGAAAACATGGATAAGACAATTCTTGTCAAAAATGAAAATATTAAATTAGAT	1620





Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAGAAGCAGTGTATGGGCCAGGTTGT	240
Qy	241	GGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTG	300
Db	241	GGTCTAGCTTGGGCTCATGCACCCATTGGATATTCTCTGAGTCTAATTTTAGGTGGTCTG	300
Qy	301	TTCTTTGCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA	360
Db	301	TTTTTTGCGAAACCTATGCGTTCCAAGGGATATGTGACTATGTTAGACCCATTCAAACAG	360
Qy	361	ATCTATGGAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTC	420
Db	361	ATCTATGGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGAGATGTTC	420
Qy	421	TGGGCTGCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGAT	480
Db	421	TGGGCTGCAGCAATTTTCTCTGCATTAGGGGCCACCATCAGCGTGATCATTGATGTGGAT	480
Qy	481	ATGCACATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGG	540
Db	481	GTGAACATATCGGTCATTGTCTCTGCACTCATTGCCATTCTTTATACCCTAGTGGGTGGG	540
Qy	541	CTCTATTCTGTGGCCTACACTGATGTCGTTTCTGCTCTTTTGCATTTTGTAGGGCTGTGG	600
Db	541	CTCTACTCTGTGGCATATACTGATGTTGTCCAGCTATTCTGCATTTTATAGGACTGTGG	600
Qy	601	ATCAGCGTCCCCTTTGCATTGTCACATCCTGCAGTCGCAGACATCGGGTTCAGTGTGTG	660
Db	601	ATCAGTGTCCCTTTTGCCCTGTCACATCCTGCAGTCACCGACATCGGATTCACAGCTGTG	660
Qy	661	CATGCCAAATACCAAAGCCGTGGCTGGGAACCTGTTGACTCATCTGAAGTCTACTCTTGG	720
Db	661	CATGCTAAATACCAGAGTCCCTGGCTGGGAACCATGAATCAGTTGAAGTCTACACCTGG	720
Qy	721	CTTGATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGG	780
Db	721	CTTGATAATTTTCTGTTATTGATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAGG	780
Qy	781	GTTCTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCCTTCCTGGCAGCTTTCGGG	840
Db	781	GTCCTCTCTTCATCCTCAGCCACCTATGCTCAGGTACTGTCCTTCCTGGCAGCTTTTGGG	840
Qy	841	TGCCCTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGG	900
Db	841	TGCCCTGGTGATGGCTCTACCCGCCATATGCATAGGAGCTATTGGAGCTTCCACAGACTGG	900
Qy	901	AACCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTA	960
Db	901	AACCAGACTGCCTACGGGTATCCAGATCCCAAGACTAAGGAGGAAGCAGACATGATTCTC	960
Qy	961	CCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCCCTGTGTACATCTCCTTCTTTGGGCTTGGTGCTGTT	1020
Qy	1021	TCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTCAGCAAGTCCATGTTTGCA	1080
Db	1021	TCAGCTGCTGTCATGTCCTCAGCTGACTCGTCCATCCTGTCGCGAGTCTATGTTTGCT	1080





Qy 1650 CATGACCCTCAGCTCAACTTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 1709  
 |||  
 Db 31295 CATGACCCTCAGCTCAACTTTTACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 31354

Qy 1710 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 1743  
 |||  
 Db 31355 AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 31388

RESULT 7

US-09-911-077A-14

; Sequence 14, Application US/09911077A  
 ; Publication No. US20030114399A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BLAKELY, RANDY D.  
 ; APPLICANT: APPARSUNDARAM, SUBRAMANIAM  
 ; APPLICANT: FERGUSON, SHAWN  
 ; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
 ; FILE REFERENCE: VBLT:008US  
 ; CURRENT APPLICATION NUMBER: US/09/911,077A  
 ; CURRENT FILING DATE: 2001-07-23  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 142299  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: Primer  
 ; FEATURE:  
 ; NAME/KEY: modified\_base  
 ; LOCATION: (1305)..(127835)  
 ; OTHER INFORMATION: N = A, C, G or T/U  
 US-09-911-077A-14

Query Match 36.2%; Score 630.8; DB 10; Length 142299;  
 Best Local Similarity 99.7%; Pred. No. 5.6e-178;  
 Matches 632; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1110 AAATGCTTCGGACAAAGAAATCGTTTGGGTATGCGAATCACAGTGTGTTGTTGGAGC 1169  
 |||  
 Db 94673 ACAGGCTTCGGACAAAGAAATCGTTTGGGTATGCGAATCACAGTGTGTTGTTGGAGC 94732

Qy 1170 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 1229  
 |||  
 Db 94733 ATCTGCAACAGCCATGGCCTTGCTGACGAAACTGTGTATGGGCTCTGGTACCTCAGTTC 94792

Qy 1230 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 1289  
 |||  
 Db 94793 TGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAAC 94852

Qy 1290 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTCCTGAGAATAACTGGAGG 1349  
 |||  
 Db 94853 CAACACCTATGGGGCCGTGGCAGGTTATGTTTCTGGCCTCTCCTGAGAATAACTGGAGG 94912

Qy 1350 GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 1409

```

Db      94913  |||||
GGAGCCATATCTGTATCTTCAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAA 94972

Qy      1410  TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAAC 1469
|||||

Db      94973  TGGTATATATAATCAGAAATTTCCATTTAAACACTTGCCATGGTTACATCATTCTTAAC 95032

Qy      1470  CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 1529
|||||

Db      95033  CAACATTTGCATCTCCTATCTAGCCAAGTATCTATTTGAAAGTGGAACCTTGCCACCTAA 95092

Qy      1530  ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 1589
|||||

Db      95093  ATTAGATGTATTTGATGCTGTTGTTGCAAGACACAGTGAAGAAAACATGGATAAGACAAT 95152

Qy      1590  TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG 1649
|||||

Db      95153  TCTTGTCAAAAATGAAAATATTAAATTAGATGAACTTGCACTTGTGAAGCCACGACAGAG 95212

Qy      1650  CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 1709
|||||

Db      95213  CATGACCCTCAGCTCAACTTTCACCAATAAAGAGGCCTTCCTTGATGTTGATTCCAGTCC 95272

Qy      1710  AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 1743
|||||

Db      95273  AGAAGGGTCTGGGACTGAAGATAATTTACAGTGA 95306

```

RESULT 8

US-10-241-784-1

; Sequence 1, Application US/10241784

; Publication No. US20040048261A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Corporation

; TITLE OF INVENTION: Invertebrate Choline Transporter Nucleic Acid,  
Polypeptides and Uses

; TITLE OF INVENTION: Thereof

; FILE REFERENCE: MO7218

; CURRENT APPLICATION NUMBER: US/10/241,784

; CURRENT FILING DATE: 2002-09-11

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1833

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1833)

; OTHER INFORMATION:

US-10-241-784-1

```

Query Match          21.6%;  Score 376.6;  DB 12;  Length 1833;
Best Local Similarity 56.0%;  Pred. No. 2e-102;
Matches 868; Conservative 0; Mismatches 624; Indels 57; Gaps 6;

```

Qy 8 TCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAA 67

Db	5	TCAATATCGCTGGCGTGGTGAGCATCGTGCTCTTCTACCTCCTGATCCTGGTCGTTGGCA	64
Qy	68	TATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATCA	127
Db	65	TTTGGGCCGGTCGCAAGAAGCAGTCCGGCAATGATTTCGGAGGAG-----GAGGTCA	115
Qy	128	TAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCTACCTGGG	187
Db	116	TGCTGGCCGGACGCTCCATCGGCCTCTTCGTGGGCATCTTCACCATGACGGCCACCTGGG	175
Qy	188	TCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAG	247
Db	176	TGGGTGGCGGCTACATCAACGGCACGGCGGAGGCTATATACACATCGGGT-----CTGG	229
Qy	248	CTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTG	307
Db	230	TGTGGTGCCAGGCTCCATTTGGATACGCTCTAAGCTTGGTATTTGGTGGCATCTTCTTTG	289
Qy	308	CAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTATG	367
Db	290	CCAATCCCATGCGCAAGCAGGGTTACATCACCATGTTGGATCCGTTGCAGGATTCTCTTTG	349
Qy	368	GAAAACGCATGGGCGGACTCCTGTTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTG	427
Db	350	GTGAGCGGATGGGAGGATTGCTCTTCTGCCGCTCTATGCGGTGAGGTCTTTTGGGCAG	409
Qy	428	CAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCACA	487
Db	410	CCGGCATCCTGGCTGCACTTGGCGCCACTCTATCGGTGATCATCGACATGGATCACCGCA	469
Qy	488	TTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATT	547
Db	470	CCTCGGTGATCCTGTCTCTGCACTCGCCATCTTCTACACACTGTTTCGGTGGACTGTACT	529
Qy	548	CTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGCG	607
Db	530	CCGTGGCGTATACGGACGTGATCCAGTTGTTCTGCATCTTCATCGGTCTGTGGATGTGCA	589
Qy	608	TCCCCTTTGCATTGTACATCCTGCAGTCGACAGATCGGGTTCACCTGCTGTGCATGCCA	667
Db	590	TTCCCTTCGCCTGGAGCAACGAGCACGTGGGCAGCCTGAGTGACCTGGAGGTGGAT----	645
Qy	668	AATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGGCTTGATA	727
Db	646	-----TGGATTGGGCACGTGGAGCCTAAAAAGCATTGGCTGTACATAGACT	691
Qy	728	GTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTCTCT	787
Db	692	ACGGCTTGCTGCTCGTCTTTGGTGGCATTCCCTGGCAGGTCTACTTCCAGCGGCAAAAC-	750
Qy	788	CTTCTTCTCAGCCACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGGTGCCTGG	847
Db	751	-----GGCAGGAAGGGCCAGCTTCTGCCTATGTTGCAGCCGCCGATGCATTT	799
Qy	848	TGATGGCCATCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAACCAGA	907

Db 800 TGATGGCCATTCCCCGGTGCTCATCGGAGCGATTGCCAAGGCTACACCTTGAACGAGA 859

Qy 908 CTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTACCAATTG 967  
 | | | | | | | | | | | | | | | | | | | | | |

Db 860 CAGATTACAAGGGACCTATCCCCTGACCGTGGACGAGACGAGCATGATTCTGCCCATGG 919

Qy 968 TTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTTTCTGCTG 1027  
 | | | | | | | | | | | | | | | | | | | | | |

Db 920 TGCTGCAGTACCTCACGCCTGACTTCGTGTCCTTCTTTGGATTGGGCGCTGTTTCCGCCG 979

Qy 1028 CTGTTATGTCATCAGCAGATTCTTCCATCTTGTGACGAAGTTCCATGTTTGCACGGAACA 1087  
 | | | | | | | | | | | | | | | | | | | | | |

Db 980 CCGTGATGTCCTCCGCCGACTCCTCGGTGCTCTCCGCCGCTCCATGTTTCGCTCGGAACG 1039

Qy 1088 TCTACCAGCTTTCCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTATGCGAA 1147  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1040 TGTACAAATTGATTTTCCGTGAGAAGGCGTCCGAGATGGAATCATTTGGGTGATGCGAG 1099

Qy 1148 TCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAAACGTGT 1207  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1100 TCGCCATCATTGTGGTGGGCATCCTGGCTACCATCATGGCCCTCACCATTCCCTCCATCT 1159

Qy 1208 ATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTTATCTTCCCCCAGCTGCTT- 1266  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1160 ACGGTTTGTGGTCCATGTGCTCGGATCTGGTCTACGTCATTCTGTTCCCGCAGCTACTGA 1219

Qy 1267 --TGTGTA CTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTGGCAGGTTATGTTTCTG 1324  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1220 TGGTGGTGCACCTCAAGAAGCACTGCAACACGTACGGCAGCCTGTCGGCATAACATTGTGG 1279

Qy 1325 GCCTCTTCCTGAGAATAACTGGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATCTTCT 1384  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1280 CCCTGGCCATCCGACTGTGCGGCGGTGAGGCCATCTTGGGACTGGCTCCATTGATCAAGT 1339

Qy 1385 ACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAATTTCCATTTAAACAC 1444  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1340 ATCCCGGCTACGACGAGGAGACCAAGG-----AGCAGATGTTCCCCTTCCGCACCA 1390

Qy 1445 TTGCCATGGTTACATCATTCTTAACCAACATTTGCATCTCCTATCTAGCCAAGTATCTAT 1504  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1391 TGGCCATGCTGCTCAGCCTGGTCACGCTCATCTCGGTCTCCTGGTGGACTAAAATGATGT 1450

Qy 1505 TTGAAAGTGGAACCTTGCCACCTAAATTAGATGTATTTGATGCTGTTGT 1553  
 | | | | | | | | | | | | | | | | | | | | | |

Db 1451 TTGAGTCCGGCAAGTTGCCGCCAGCTACGACTACTTCCGCTGTGTGGT 1499

RESULT 9

US-09-911-077A-7

; Sequence 7, Application US/09911077A

; Publication No. US20030114399A1

; GENERAL INFORMATION:

; APPLICANT: BLAKELY, RANDY D.

; APPLICANT: APPARSUNDARAM, SUBRAMANIAM

; APPLICANT: FERGUSON, SHAWN

; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA

; FILE REFERENCE: VBLT:008US



; CURRENT APPLICATION NUMBER: US/09/911,077A  
; CURRENT FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1985  
; TYPE: DNA  
; ORGANISM: *Caenorhabditis elegans*  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (9)..(1739)  
US-09-911-077A-7

Query Match 20.9%; Score 363.8; DB 10; Length 1985;  
Best Local Similarity 55.1%; Pred. No. 1.6e-98;  
Matches 862; Conservative 0; Mismatches 637; Indels 66; Gaps 5;

```
Qy      19 GGA CTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTGGTTGGAATATGGGCTGCC 78
      || | | || | | | | | | | | | | | | | | | | | | | | | | | | |
Db      24 GGTATCGTGGCCATTGTGTTCTTCTACGTGCTCATTCTTGTCGTTGGAATATGGGCGGGT 83

Qy      79 TGGAGAACCAAAA-----ACAGTGGCAGCGCAGAAGAGCGCAGCGAAGCCATC 126
      | | | | || | | | | | | | | | | | | | | | | | | | | | | |
Db      84 AGAAAATCGAAAAGTTCAAAAAGAGCTTGAATCAGAAGCCGGCGCGGCGACGGAAGAGGTG 143

Qy     127 ATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTTACCATGACAGCTACCTGG 186
      || | | || | | | | | | | | | | | | | | | | | | | | | | |
Db     144 ATGTTAGCTGGGAGAAACATCGAACTCTTGTCGGAATTTTCACAATGACTGCCACGTGG 203

Qy     187 GTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTA 246
      || | | || | | | | | | | | | | | | | | | | | | | | | | |
Db     204 GTTGGCGGTGCTTATATCAATGGAACCGCCGAGGCTCTGTATAATGGAGGT-----CTC 257

Qy     247 GCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTT 306
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     258 CTTGGATGTCAGGCTCCAGTTGGATATGCAATTTCCCTTGTTATGGGAGGACTACTTTTC 317

Qy     307 GCAAAACCTATGCGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAAATCTAT 366
      || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     318 GCAAAGAAAATGCGAGAAGAAGGATATATTACAATGCTCGATCCTTTTCAGCACAAATAT 377

Qy     367 GGAAAACGCATGGGCGGACTCCTGTTTATTCCTGCACTGATGGGAGAAATGTTCTGGGCT 426
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     378 GGCCAACGAATCGGTGGCTTGATGTATGTCCAGCACTTCTTGGTGAAACATTCTGGACA 437

Qy     427 GCAGCAATTTTCTCTGCTTTGGGAGCCACCATCAGCGTGATCATCGATGTGGATATGCAC 486
      || || | | | | | | | | | | | | | | | | | | | | | | | | |
Db     438 GCAGCCATTCTTTCGGCACTTGGTGCAACACTGTCGGTAATTCTTGAATCGACATGAAT 497

Qy     487 ATTTCTGTCATCATCTCTGCACTCATTGCCACTCTGTACACACTGGTGGGAGGGCTCTAT 546
      || || | | | | | | | | | | | | | | | | | | | | | | | | |
Db     498 GCATCAGTGACCCTGTGGCCTGTATTGCCGTATTCTACACATTCACCGGTGGATACTAT 557

Qy     547 TCTGTGGCCTACACTGATGTCGTTTCAGCTCTTTTGATTTTGTAGGGCTGTGGATCAGC 606
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     558 GCAGTCGCGTACACTGACGTCGTTCAACTATTTTGATTTTCGTCGTTTGTGGGTTTGC 617
```

Qy	607	GTCCCTTTTGCATTGTCACATCCTGCAGTCGACATCGGGTTCACTGTGTCATGCC	666
Db	618	GTGCCGGCGGCTATGGTGCATGATGGTGCGAAGGATATTTCCAGGAATGCAG-----	669
Qy	667	AAATACCAAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGGCTTGAT	726
Db	670	-----GCGACTGGATTGGAGAGATTGGAGGATTCAAAGAACATCTCTCTGGATTGAT	722
Qy	727	AGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTC	786
Db	723	TGCATGCTTCTCCTTGTCTTTGGAGGAATCCCATGGCAAGTGTACTTCCAAAGAGTTCTC	782
Qy	787	TCTTCTTCTCTCAGCCACCTATGCTCAAGTGCTGTCTTCTGGCAGCTTTCGGGTGCCTG	846
Db	783	TCCTCAAAAAGTCTCATGGAGCACAGACGTTGTCTGTTTGTGGCGGGCGTCGGATGCATT	842
Qy	847	GTGATGGCCATCCCAGCCATACTCATTTGGGGCCATTGGAGCATCAACAGACTGGAACCAG	906
Db	843	CTCATGGCGATTCCACCAGCGTTGATCGGTGCAATTGCCAGGAACACAGACTGGAGAATG	902
Qy	907	ACTGCATATGGGCTTCC-----AGATCCAAGACTACAGAAGAGGCA	948
Db	903	ACTGATTATTCCCCATGGAACAATGGAACTAAGGTGCAATCGATTCCACCGGATAAGAGA	962
Qy	949	GACATGATTTTACCAATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGT	1008
Db	963	AACATGGTGGTCCCGTTGGTATTCCAGTATCTTACGCCAAGATGGGTGCCTTTATTGGA	1022
Qy	1009	CTTGGTGCAAGTTTCTGCTGCTGTTATGTCATCAGCAGATTCTTCCATCTTGTGCAAGT	1068
Db	1023	CTCGGCGCAGTGTGCGCTGCTGTAATGTCATCTGCAGATTCATCTGTACTATCAGCAGCA	1082
Qy	1069	TCCATGTTTGCACGGAACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAA	1128
Db	1083	TCAATGTTTGCTCACAACATCTGGAAGCTCACAATTCGCCCTCACGCGTCTGAAAAAGAA	1142
Qy	1129	ATCGTTTGGGTATGCGAATCACAGTGTGTTGTGTTTGGAGCATCTGCAACAGCCATGGCC	1188
Db	1143	GTGATAATTGTGATGAGAATAGCCATCATCTGTGTTGGTATCATGGCAACCATCATGGCA	1202
Qy	1189	TTGCTGACGAAAAGTGTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTT	1248
Db	1203	CTTACCATCAATCCATCTATGGGCTTGGTATCTTTGTGCAGATTGGTCTACGTGATA	1262
Qy	1249	ATCTTCCCCAGCTGCTTTGTGTACTCTTTGTTAAGGGAACCAACACCTATGGGGCCGTG	1308
Db	1263	CTCTTCCCTCAACTATTATGTGTTGTATATATGCCACGTAGCAATACGTATGGCTCATTG	1322
Qy	1309	GCAGGTTATGTTTCTGGCCTCTTCTTGAGAATAACTGGAGGGGAGCCATATCTGTATCTT	1368
Db	1323	GCTGGCTATGCAGTCGGTCTTGTGCTCCGTTTGATTGGAGGCGAGCCACTTGTATCGCTG	1382
Qy	1369	CAGCCCTTGATCTTCTACCCTGGCTATTACCCTGATGATAATGGTATATATAATCAGAAA	1428
Db	1383	CCAGCGTTCTTCCATTATCCAATGTATACGGATGGGG-----TACAGTAT	1427
Qy	1429	TTTCCATTTAAACACTTGCCATGGTTACATCATTTCTTAACCAACATTTGCATCTCCTAT	1488

```

      || |||| | ||| ||| ||| | ||| | || | | ||
Db      1428 TTCCCATTCAGGACAACTGCTATGTTATCTTCAATGGCTACTATCTACATTGTATCAATA 1487
Qy      1489 CTAGCCAAGTATCTATTTGAAAGTGGAAACCTTGCCACCTAAATTAGATGTATTTGATGCT 1548
      | | | | | | | | | | | | | | | | | | | |
Db      1488 CAATCGGAGAAGCTGTTCAAATCGGGACGTTTGTCTCCGGAGTGGGACGTAATGGGTTGT 1547
Qy      1549 GTTGT 1553
      || ||
Db      1548 GTAGT 1552

```

RESULT 10

US-09-974-300-501

```

; Sequence 501, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-501

```

```

Query Match          13.9%; Score 242.6; DB 9; Length 1461;
Best Local Similarity 52.4%; Pred. No. 5.9e-62;
Matches 663; Conservative 0; Mismatches 554; Indels 48; Gaps 4;

```

```

Qy      151 TTATTGGTTGGTGGATTTACCATGACAGCTACCTGGGTCGGAGGAGGGTATATCAATGGC 210
      || || || || | || ||| | || || ||||| || || ||||| || ||
Db      125 TTTTTCGTCGGAATGGTGACGATGGCCGCAACATGGGTCGGCGGCGGATATATTAACGGA 184
Qy      211 ACAGCTGAAGCAGTTTATGTACCAGGTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGA 270
      || || ||| | ||| | | |||| | ||| || || || ||
Db      185 ACGGCCGAATCGACTTACA-----GCGACGGCCTCATCTGGGCCCAAGCGCCTTGGGGC 238
Qy      271 TATTCTCTTAGTCTGATTTTAGGTGGCCTGTTCTTTGCAAACCTATGCGTTCAAAGGGG 330
      || | | || ||||| | || || | |||| || | | |||| |
Db      239 TACGCATTGAGCCTGATTATCGGCGGTATTTTCTTCGCCAGAAAAATGCGCCGTCAACAA 298
Qy      331 TATGTGACCATGTTAGACCCGTTTCAGCAAATCTATGGAAAACGCATGGGCGGACTCCTG 390
      | | ||||| | || || | | || || || || || || || | |
Db      299 TTTATGACCATTATCGATCCTCTCGAACAGCGCTTCGGTAAGCGGATGGCCGGCGTTTAA 358
Qy      391 TTTATTCTGCACTGATGGGAGAAATGTTCTGGGCTGCAGCAATTTTCTCTGCTTTGGGA 450

```

Db 359 TATATACCGGCGCTGTTAGGAGAATTGTTTTGGAGCGCCGCGATCTTAACGGCATTGGGC 418  
 Qy 451 GCCACCATCAGCGTGATCATCGATGTGGATATGCACATTTCTGTCATCATCTCTGCACTC 510  
 Db 419 ACGACTTTCGGAATGATTCTGAATATCGATTTCCAAACGTCGATTATTCTTTTCGGCGATG 478  
 Qy 511 ATTGCCACTCTGTACACACTGGTGGGAGGGCTCTATTCTGTGGCCTACACTGATGTCGTT 570  
 Db 479 ATCGCCATCGCTTATACGGTGGCCGGAGGCATGTGGGCAGTTGCTTTCACAGATGTCTTT 538  
 Qy 571 CAGCTCTTTTGCATTTTGTAGGGCTGTGGATCAGCGTCCCCTTTGCATTGTCACATCCT 630  
 Db 539 CAAATGATTGTCATTTTGTCTCGGGCTGTTTTTAGTCGTCCCATTTGTATTGTCGAATGTC 598  
 Qy 631 GCAG-----TCGCAGACATCGGGTTCAGTGTGCTGTGCATGCCAA 668  
 Db 599 GGTGCTCTCGATAGCGTATGGGCAAATTACAGGCACGATTTTCGGCAGCAGCGCCAATCTG 658  
 Qy 669 ATACCAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCT-----TGGCTT 723  
 Db 659 CTTCCGCCGTGGACGGCTGGAAAAACCCGGATTGGGGAAACCTGTTTTGGAACGGTGG 718  
 Qy 724 GATAGTTTTCTGTTGTTGATGCTGGGTGGAATCCCATGGCAAGCATACTTTTCAGAGGGTT 783  
 Db 719 GATAATGCGCTCCTCTTAATTTTCGGAGGTATCGCATGGCAGGTGTACTTTTCAGCGCGTT 778  
 Qy 784 CTCTCTTCTTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCCTGGCAGCTTTTCGGGTGC 843  
 Db 779 CTTTCGGCAAAATCGGAAAGCGCCGCCATGTGGCAGTCGATAATTGCCGGAGTGATCTGC 838  
 Qy 844 CTGGTGATGGCCATCCCAGCCATACTCATTGGGGCCATTGGAGCATCAACAGACTGGAAC 903  
 Db 839 ATCATTGCCGCCATTCCGTGCGTAATCATCGGAGCTGCCGGAACAGTACCGATTGGAGC 898  
 Qy 904 CAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAGAGGCAGACATGATTTTACCA 963  
 Db 899 CTGTTTCGGAGCGAGCGCTCCGGATAACCCGGCG-----ATGATTTTGCCG 943  
 Qy 964 ATTGTTCTGCAGTATCTCTGCCCTGTGTATATTTCTTTCTTTGGTCTTGGTGCAGTTTCT 1023  
 Db 944 CAAACGCTTGCGTATTTGACGCCAGGAATCATCGCAGGCCTCGGCTTGGGTGCAATCGCA 1003  
 Qy 1024 GCTGCTGTTATGTCATCAGCAGATTCTTCCATCTGTGTCAGCAAGTTCCATGTTTGACGG 1083  
 Db 1004 GCAGCCGTCATGTCAAGCATGGACTCATCGATTCTATCGGCATCATCAATGGCCGCATGG 1063  
 Qy 1084 AACATCTACCAGCTTTCCTTCAGACAAAATGCTTCGGACAAAGAAATCGTTTGGGTATG 1143  
 Db 1064 AATATTTACCGTCCGCTCATCAAGCCGAAGGCCACCCAAAAACAGCTGCAAAAAGTCGTC 1123  
 Qy 1144 CGAATCACAGTGTGTTGTGTTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAAACT 1203  
 Db 1124 AAACGCTCAATCATTTTGTTCGGCGCGGGAGCAGCGGTGTCGCGCTGAATGTCAAAAGC 1183  
 Qy 1204 GTGTATGGGCTCTGGTACCTCAGTTCTGACCTTGTGTTTACATCGTTATCTTCCCCAGCTG 1263



Db 94404 ACGT 94401

RESULT 12

US-09-911-077A-14/c

; Sequence 14, Application US/09911077A  
; Publication No. US20030114399A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARSUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/09/911,077A  
; CURRENT FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 142299  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (1305)..(127835)  
; OTHER INFORMATION: N = A, C, G or T/U  
US-09-911-077A-14

Query Match 10.4%; Score 180.8; DB 10; Length 142299;  
Best Local Similarity 98.9%; Pred. No. 7.8e-42;  
Matches 182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 60  
|||||  
Db 64222 ATGGCTTTCCATGTGGAAGGACTGATAGCTATCATCGTGTCTACCTTCTAATTTTGCTG 64163  
  
Qy 61 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 120  
|||||  
Db 64162 GTTGAATATGGGCTGCCTGGAGAACCAAAACAGTGGCAGCGCAGAAGAGCGCAGCGAA 64103  
  
Qy 121 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGCT 180  
|||||  
Db 64102 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATTGGTTGGTGGATTACCATGACAGGT 64043  
  
Qy 181 ACCT 184  
|||  
Db 64042 ACGT 64039

RESULT 13

US-09-864-761-1838

; Sequence 1838, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 1838  
; LENGTH: 455  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC009963.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.97  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
US-09-864-761-1838

Query Match 8.9%; Score 155; DB 9; Length 455;  
Best Local Similarity 100.0%; Pred. No. 8.1e-36;  
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 GATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTCTCTTCTCCTCAGC 800  
|||||  
Db 266 GATGCTGGGTGGAATCCCATGGCAAGCATACTTTCAGAGGGTTCTCTCTTCTCCTCAGC 325  
  
Qy 801 CACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGGTGCCTGGTGATGGCCATCCC 860  
|||||  
Db 326 CACCTATGCTCAAGTGCTGTCCTTCTGGCAGCTTTCGGGTGCCTGGTGATGGCCATCCC 385  
  
Qy 861 AGCCATACTCATTGGGGCCATTGGAGCATCAACAG 895  
|||||  
Db 386 AGCCATACTCATTGGGGCCATTGGAGCATCAACAG 420

RESULT 14

US-10-027-632-120553/c

; Sequence 120553, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 120553  
; LENGTH: 943  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-120553

Query Match 6.8%; Score 118.6; DB 15; Length 943;



Best Local Similarity 99.2%; Pred. No. 1.4e-24;  
Matches 118; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 176 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 235  
|||||  
Db 589 CAGCTACCTGGGTCGGAGGAGGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACCAG 530  
Qy 236 GTTATGGCCTAGCTTGGGCTCAGGCACCAATTGGATATTCTCTTAGTCTGATTTTAGGT 294  
|||||:|||||  
Db 529 GTTATGGCCTAGCTTGGGCTCAGGCACCARTGGATATTCTCTTAGTCTGATTTTAGGT 471

RESULT 15

US-09-864-761-18589

; Sequence 18589, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES

USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18589
; LENGTH: 96
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009963.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: NT HIT: gill1141884, EVALUE 5.00e-33
US-09-864-761-18589

```

```

Query Match          4.1%; Score 72; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      824 TCCTGGCAGCTTTCGGGTGCCTGGTGGTATGGCCATCCCAGCCATACTCATTGGGGCCATTG 883
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCCTGGCAGCTTTCGGGTGCCTGGTGGTATGGCCATCCCAGCCATACTCATTGGGGCCATTG 60

Qy      884 GAGCATCAACAG 895
        |||||||||||
Db      61 GAGCATCAACAG 72

```

```

Search completed: March 22, 2004, 15:30:20
Job time : 650 secs

```